

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:15:50 ; Search time 172 Seconds
(without alignments)
2188.245 Million cell updates/sec

Title: US-10-038-972A-13

Perfect score: 3994

Sequence: 1 MAADGYLPDWMLEDTLSEGR.....TNGVYSEPRPTGTRYLRNL 735

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_eprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3994	100.0	735	2	O56652	adeno-associ
2	3973	99.5	735	2	Q6JC41	adeno-associ
3	3972	99.4	735	2	Q6JC06	adeno-associ
4	3969	99.4	735	2	Q6JC20	adeno-associ
5	3968	99.3	735	2	Q6JC09	adeno-associ
6	3967	99.3	735	2	Q6JC11	adeno-associ
7	3961	99.2	735	2	Q6JC01	adeno-associ
8	3961	99.2	735	2	Q6JC21	adeno-associ
9	3955	99.0	735	2	Q6JC25	adeno-associ
10	3954	99.0	735	2	Q6JC25	adeno-associ
11	3953	99.0	735	2	Q6JC29	adeno-associ
12	3946	98.8	735	2	Q6JC07	adeno-associ
13	3945	98.8	735	2	Q6JC07	adeno-associ
14	3941	98.7	735	2	Q6JC06	adeno-associ
15	3940	98.6	735	2	Q6JC25	adeno-associ
16	3939	98.6	735	2	Q6JC26	adeno-associ
17	3937	98.6	735	2	Q6JC06	adeno-associ
18	3934	98.5	735	2	Q6JC07	adeno-associ
19	3933	98.5	735	2	Q6JC29	adeno-associ
20	3932	98.4	735	2	Q6JC29	adeno-associ
21	3931	98.4	735	2	Q6JC24	adeno-associ
22	3928.5	98.4	734	2	Q6JC00	adeno-associ
23	3927	98.3	735	2	Q6JC00	adeno-associ
24	3923	98.2	735	2	Q6JC00	adeno-associ
25	3921	98.1	735	2	Q6JC07	adeno-associ
26	3917	98.1	735	2	Q6JC31	adeno-associ
27	3916	98.0	735	2	Q6JC27	adeno-associ
28	3916	98.0	735	2	Q6JC30	adeno-associ
29	3904	97.7	735	2	Q6JC33	adeno-associ
30	3725	93.3	735	2	Q6JC27	adeno-associ
31	3711	92.9	735	2	Q6JC34	adeno-associ

32	3709.5	92.9	734	2	Q6JC02	Q6JC02 adeno-associ
33	3706	92.8	735	2	Q6JC17	Q6JC17 adeno-associ
34	3704.5	92.8	734	2	Q6JC04	Q6JC04 adeno-associ
35	3704	92.7	735	2	Q6JB21	Q6JB21 adeno-associ
36	3704	92.7	735	2	Q6JC42	Q6JC42 adeno-associ
37	3703	92.7	735	2	Q6JB26	Q6JB26 adeno-associ
38	3701	92.7	735	2	Q6JC44	Q6JC44 adeno-associ
39	3700	92.6	735	2	Q6JC28	Q6JC28 adeno-associ
40	3690.5	92.4	734	2	Q6JC03	Q6JC03 adeno-associ
41	3688	92.3	735	2	Q6JC38	Q6JC38 adeno-associ
42	3685	92.3	735	2	Q6JC36	Q6JC36 adeno-associ
43	3683	92.2	735	2	Q6JC08	Q6JC08 adeno-associ
44	3683	92.2	735	2	Q6JC43	Q6JC43 adeno-associ
45	3676	92.0	735	2	Q6JB20	Q6JB20 adeno-associ

ALIGNMENTS

RESULT 1

```
O56652 ID O56652 PRELIMINARY; PRT; 735 AA.
AC O56652;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE Major coat protein VP1.
DE Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_taxid=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid
RT proteins affect viral infectivity: lack of an RGD integrin-binding
RT motif."
RL J. Gen. Virol. 75:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF043303; AAC03780.1;
DR GO; GO:0015928; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
SQ SEQUENCE 735 AA; 81944 MW; 980BEEF46908390B CRC64;
```

Query Match	100.0%;	Score	3994;	DB	2;	Length	735;
Best Local Similarity	100.0%;	Pred. No.	1.4e-242;				
Matches	735;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAADGYLPDWMLEDTLSEGIQWVKLKGPPPPKPAERHKDSDRGVLVLPYKYLGPENGLD	60				
Db	1	MAADGYLPDWMLEDTLSEGIQWVKLKGPPPPKPAERHKDSDRGVLVLPYKYLGPENGLD	60				
Qy	61	KGEVNEADAALAEHDKAYDRQLSDGNPNLYKYNHADAERFKEDTSFGNIGRAVFQ	120				
Db	61	KGEVNEADAALAEHDKAYDRQLSDGNPNLYKYNHADAERFKEDTSFGNIGRAVFQ	120				
Qy	121	AKKRVLPGLVPEPVKTAPGKRPVHSVPVSSSGTGKAGQOPARKLNFGQTGDAD	180				
Db	121	AKKRVLPGLVPEPVKTAPGKRPVHSVPVSSSGTGKAGQOPARKLNFGQTGDAD	180				
Qy	181	SVDPDQPLGGPPAPSPGLGTNTMTATGSGAPMADNNEGADVGNSSGNHCDSTWMDRVI	240				
Db	181	SVDPDQPLGGPPAPSPGLGTNTMTATGSGAPMADNNEGADVGNSSGNHCDSTWMDRVI	240				
Qy	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCFSPRDWQRLI	300				

```
Db      241  TTSTRWALPTNNHLYKQISQSGASNDNHFGYSTPWGYDFNRFCHFSRDMQRLI 300
Qy      301  NNNWGRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Db      301  NNNWGRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Qy      361  CLPPPPADVFMPVQGYLTLLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEDVVF 420
Db      361  CLPPPPADVFMPVQGYLTLLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEDVVF 420
Qy      421  HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db      421  HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Qy      481  PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
Db      481  PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
Qy      541  IFKGQSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRNQAATADVNTQGV 600
Db      541  IFKGQSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRNQAATADVNTQGV 600
Qy      601  LPMQWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db      601  LPMQWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy      661  FSAAKFASFTQYSTQGVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVDTNGVY 720
Db      661  FSAAKFASFTQYSTQGVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVDTNGVY 720
Qy      721  SEPRPIGTRYLTRNL 735
Db      721  SEPRPIGTRYLTRNL 735

RESULT 2
Q6JC41  PRELIMINARY; PRT; 735 AA.
ID Q6JC41
AC Q6JC41;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530578; AAS99263.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81748 MW; 1555801B683FE750 CRC64;

Query Match 99.58; Score 3973; DB 2; Length 735;
Best Local Similarity 99.58; Pred. No. 3e-241;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSGIRQWKKLKPFPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSGIRQWKKLKPFPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGFPVNEADAAAEHDKAYDRQLDSGNDPNLYKYNHADAEEQERLKEDTSFGNLRGAVFQ 120
```

```
Db      61  KGFPVNEADAAAEHDKAYDRQLDSGNDPNLYKYNHADAEEQERLKEDTSFGNLRGAVFQ 120
Qy      121  AKKRVLPLGLVEBPKTAPCKKRPVHSVPBESSCTGKAGQOPARKRLNFGGTGDAD 180
Db      121  AKKRVLPLGLVEBPKTAPCKKRPVHSVPBESSCTGKAGQOPARKRLNFGGTGDAD 180
Qy      181  SVDPDQPLGQCPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNHCHDSTWMDRVI 240
Db      181  SVDPDQPLGQCPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNHCHDSTWMDRVI 240
Qy      241  TTSTRWALPTNNHLYKQISQSGASNDNHFGYSTPWGYDFNRFCHFSRDMQRLI 300
Db      241  TTSTRWALPTNNHLYKQISQSGASNDNHFGYSTPWGYDFNRFCHFSRDMQRLI 300
Qy      301  NNNWGRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Db      301  NNNWGRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Qy      361  CLPPPPADVFMPVQGYLTLLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEDVVF 420
Db      361  CLPPPPADVFMPVQGYLTLLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEDVVF 420
Qy      421  HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db      421  HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Qy      481  PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
Db      481  PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
Qy      541  IFKGQSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRNQAATADVNTQGV 600
Db      541  IFKGQSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRNQAATADVNTQGV 600
Qy      601  LPMQWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db      601  LPMQWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy      661  FSAAKFASFTQYSTQGVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVDTNGVY 720
Db      661  FSAAKFASFTQYSTQGVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVDTNGVY 720
Qy      721  SEPRPIGTRYLTRNL 735
Db      721  SEPRPIGTRYLTRNL 735

RESULT 3
Q6JC06  PRELIMINARY; PRT; 735 AA.
ID Q6JC06
AC Q6JC06;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530613; AAS99298.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81786 MW; 802F42CBFCFEFIC1 CRC64;
```

Query Match		99.4%; Score 3972; DB 2; Length 735;
Best Local Similarity		99.5%; Pred. No. 3.4e-241;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;		
Qy	1	MAADGYLPDWMLEDTLSGIRQWKKLPGPPPPKPAERHKDDSRGLVLPVGYKYLGPFGNGLD 60
Db	1	MAADGYLPDWMLEDTLSGIRQWKKLPGPPPPKPAERHKDDSRGLVLPVGYKYLGPFGNGLD 60
Qy	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPNLYKYNHADAEEFOERLKEDTSFGGNLGRAVFQ 120
Db	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPNLYKYNHADAEEFOERLKEDTSFGGNLGRAVFQ 120
Qy	121	AKKRVLEPLGLVBEVPKTAPOKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQGTGDAD 180
Db	121	AKKRVLEPLGLVBEVPKTAPOKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQGTGDAD 180
Qy	181	SVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVNSGNNWHCDSTWMDRVI 240
Db	181	SVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVNSGNNWHCDSTWMDRVI 240
Qy	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFPSPRDWQRLI 300
Db	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFPSPRDWQRLI 300
Qy	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSYQLPYVLGSAHQG 360
Db	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSYQLPYVLGSAHQG 360
Qy	361	CLPPFPADVFMVPOGYLTLNNGSOAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420
Db	361	CLPPFPADVFMVPOGYLTLNNGSOAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420
Qy	421	HSSVAHSQSIDLRLMNPILIDQLYLSTRTNTPSGTTTOSRLQFSQAGASDIRDQSRNWLPG 480
Db	421	HSGVAHSQSIDLRLMNPILIDQLYLSTRTNTPSGTTTOSRLQFSQAGASDIRDQSRNWLPG 480
Qy	481	PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQSGVL 540
Db	481	PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDNEKFFPQSGVL 540
Qy	541	IFGQSGSEKTNVDIEKWMITDEBEIRTTNPVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
Db	541	IFGQSGSEKTNVDIEKWMITDEBEIRTTNPVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
Qy	601	LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVANPSTT 660
Db	601	LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVANPSTT 660
Qy	661	FSAAKFASFITQYSTQGVSVIEIWELOKENSKRWNPEIQYTSNYSKSVNVDFTVDTNGVY 720
Db	661	FSAAKFASFITQYSTQGVSVIEIWELOKENSKRWNPEIQYTSNYSKSVNVDFTVDTNGVY 720
Qy	721	SEPRPIGTRVLTNRNL 735
Db	721	SEPRPIGTRVLTNRNL 735
RESULT 4		
ID	Q6JC20	
AC	Q6JC20;	PRELIMINARY; PRT; 735 AA.
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Capsid protein VP1.	
GN	Name=cap;	
OS	Adeno-associated virus.	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.	
OX	NCBI_TaxID=272636;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;	
Query Match		99.4%; Score 3969; DB 2; Length 735;
Best Local Similarity		99.6%; Pred. No. 5.3e-241;
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;		
Qy	1	MAADGYLPDWMLEDTLSGIRQWKKLPGPPPPKPAERHKDDSRGLVLPVGYKYLGPFGNGLD 60
Db	1	MAADGYLPDWMLEDTLSGIRQWKKLPGPPPPKPAERHKDDSRGLVLPVGYKYLGPFGNGLD 60
Qy	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPNLYKYNHADAEEFOERLKEDTSFGGNLGRAVFQ 120
Db	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPNLYKYNHADAEEFOERLKEDTSFGGNLGRAVFQ 120
Qy	121	AKKRVLEPLGLVBEVPKTAPOKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQGTGDAD 180
Db	121	AKKRVLEPLGLVBEVPKTAPOKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQGTGDAD 180
Qy	181	SVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVNSGNNWHCDSTWMDRVI 240
Db	181	SVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVNSGNNWHCDSTWMDRVI 240
Qy	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFPSPRDWQRLI 300
Db	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFPSPRDWQRLI 300
Qy	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSYQLPYVLGSAHQG 360
Db	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSYQLPYVLGSAHQG 360
Qy	361	CLPPFPADVFMVPOGYLTLNNGSOAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420
Db	361	CLPPFPADVFMVPOGYLTLNNGSOAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420
Qy	421	HSSVAHSQSIDLRLMNPILIDQLYLSTRTNTPSGTTTOSRLQFSQAGASDIRDQSRNWLPG 480
Db	421	HSGVAHSQSIDLRLMNPILIDQLYLSTRTNTPSGTTTOSRLQFSQAGASDIRDQSRNWLPG 480
Qy	481	PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQSGVL 540
Db	481	PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDNEKFFPQSGVL 540
Qy	541	IFGQSGSEKTNVDIEKWMITDEBEIRTTNPVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
Db	541	IFGQSGSEKTNVDIEKWMITDEBEIRTTNPVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
Qy	601	LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVANPSTT 660
Db	601	LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVANPSTT 660
Qy	661	FSAAKFASFITQYSTQGVSVIEIWELOKENSKRWNPEIQYTSNYSKSVNVDFTVDTNGVY 720
Db	661	FSAAKFASFITQYSTQGVSVIEIWELOKENSKRWNPEIQYTSNYSKSVNVDFTVDTNGVY 720
Qy	721	SEPRPIGTRVLTNRNL 735
Db	721	SEPRPIGTRVLTNRNL 735
RESULT 5		
ID	Q6JC09	
AC	Q6JC09;	PRELIMINARY; PRT; 735 AA.

```

05-JUL-2004 (TrEMBLrel. 27, Created)
05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]_
SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530610; AAS99295.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat_1.
DR SEQUENCE 735 AA; 81817 MW; 496DF18E95CABB13 CRC64;
SQ
Query Match          99.3%; Score 3968; DB 2; Length 735;
Best Local Similarity 99.3%; Pred. No. 6.1e-241;
Matches 730; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAADGYLPDWEEDTLSEGIROWKLPKPGPPPKPAERHKDDSGGLVPGYKYLGPNGLD 60
DB 1 MAADGYLPDWEEDTLSEGIROWKLPKPGPPPKPAERHKDDSGGLVPGYKYLGPNGLD 60
QY 61 KGPVNEADAAALEHDKAYDRQLDSDGNPVLKYNHADAEFOERLKEDTSFGNLAGRAVQ 120
DB 61 KGPVNEADAAALEHDKAYDRQLDSDGNPVLKYNHADAEFOERLKEDTSFGNLAGRAVQ 120
QY 121 AKKRVLEPLGLVEEPVKTAPGKRVPVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKRVPVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
QY 181 SVDPDQPLGQPPAAPSGLGTNTMATSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240
DB 181 SVDPDQPLGQPPAAPSGLGTNTMATSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240
QY 241 TTSTRWALPTYNHLYKQISSQSGASNDNHYFGYSTPMGYDFNRFHCHFSPRDWQRLI 300
DB 241 TTSTRWALPTYNHLYKQISSQSGASNDNHYFGYSTPMGYDFNRFHCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
DB 301 NNNWGFPRKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPFPADVPMVPOQYGLTLNNGSQAVGRSSFYCLEYFPQSMLRTGNNTFFSYTPEDVFP 420
DB 361 CLPPFPADVPMVPOQYGLTLNNGSQAVGRSSFYCLEYFPQSMLRTGNNTFFSYTPEDVFP 420
QY 421 HSSYAHQSLSRLMNLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
DB 421 HSSYAHQSLSRLMNLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKEKFFPOSQV 540
DB 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKEKFFPOSQV 540
QY 541 IFKGQSEKTNVDIEKVMITDEEIEITNPVATEQYGSVSTNLQRNQAATADVNTQGV 600
DB 541 IFKGQSEKTNVDIEKVMITDEEIEITNPVATEQYGSVSTNLQRNQAATADVNTQGV 600
QY 601 LPMQWQDRDVLQGPIMAKIPHTDGHFSPPLMGGFGLKHPPPQILIKNTVPANPSTT 660
DB 601 LPMQWQDRDVLQGPIMAKIPHTDGHFSPPLMGGFGLKHPPPQILIKNTVPANPSTT 660
QY 661 FSAAKFASFITQYSTQVSVEIWELOKENSKRWNPEIQYTSYNKSVNVDFVTDTNGVY 720
DB 661 FSAAKFASFITQYSTQVSVEIWELOKENSKRWNPEIQYTSYNKSVNVDFVTDTNGVY 720

```

```

661 FSAAKFASFITQYSTQVSVEIWELOKENSKRWNPEIQYTSYNKSVNVDFVTDTNGVY 720
721 SEPRPIGTRYLTRNL 735
721 SEPRPIGTRYLTRNL 735
RESULT 6
Q6JC11 PRELIMINARY; PRT; 735 AA.
ID Q6JC11
AC Q6JC11;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]_
SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530608; AAS99293.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat_1.
DR SEQUENCE 735 AA; 81797 MW; 7A6CC0B6C0C78F0D CRC64;
SQ
Query Match          99.3%; Score 3967; DB 2; Length 735;
Best Local Similarity 99.3%; Pred. No. 7e-241;
Matches 730; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAADGYLPDWEEDTLSEGIROWKLPKPGPPPKPAERHKDDSGGLVPGYKYLGPNGLD 60
DB 1 MAADGYLPDWEEDTLSEGIROWKLPKPGPPPKPAERHKDDSGGLVPGYKYLGPNGLD 60
QY 61 KGPVNEADAAALEHDKAYDRQLDSDGNPVLKYNHADAEFOERLKEDTSFGNLAGRAVQ 120
DB 61 KGPVNEADAAALEHDKAYDRQLDSDGNPVLKYNHADAEFOERLKEDTSFGNLAGRAVQ 120
QY 121 AKKRVLEPLGLVEEPVKTAPGKRVPVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKRVPVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
QY 181 SVDPDQPLGQPPAAPSGLGTNTMATSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240
DB 181 SVDPDQPLGQPPAAPSGLGTNTMATSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240
QY 241 TTSTRWALPTYNHLYKQISSQSGASNDNHYFGYSTPMGYDFNRFHCHFSPRDWQRLI 300
DB 241 TTSTRWALPTYNHLYKQISSQSGASNDNHYFGYSTPMGYDFNRFHCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
DB 301 NNNWGFPRKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPFPADVPMVPOQYGLTLNNGSQAVGRSSFYCLEYFPQSMLRTGNNTFFSYTPEDVFP 420
DB 361 CLPPFPADVPMVPOQYGLTLNNGSQAVGRSSFYCLEYFPQSMLRTGNNTFFSYTPEDVFP 420
QY 421 HSSYAHQSLSRLMNLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
DB 421 HSSYAHQSLSRLMNLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKEKFFPOSQV 540
DB 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKEKFFPOSQV 540

```

```
Qy 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNQAATAADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNQAATAADVNTQGV 600
Qy 601 LPGMVQDRDVLVLOGPIWAKIPHTDGHFHPSPLMGGFGLKHPPOQILIKNTPVPANPSTT 660
Db 601 LPGMVQDRDVLVLOGPIWAKIPHTDGHFHPSPLMGGFGLKHPPOQILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNTYKNSVNVDFTTVDNTGVY 720
Db 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNTYKNSVNVDFTTVDNTGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 7
Q6JC01
ID Q6JC01 PRELIMINARY; PRT; 735 AA.
AC Q6JC01;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530618; AAS99303.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81834 MW; ACF51A69B66B582D CRC64;

Query Match 99.2%; Score 3961; DB 2; Length 735;
Best Local Similarity 98.9%; Pred. No. 1.7e-240;
Matches 721; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWMLEDTLSGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Db 1 MAADGYLPDWMLEDTLSGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Qy 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHDAEAEFOERLKEDTSFGNGLGRAVFQ 120
Db 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHDAEAEFOERLKEDTSFGNGLGRAVFQ 120
Qy 121 AKKRVLEPLGLVEEPVKTAPEGKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPEGKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAPSPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVV 240
Db 181 SVDPDQPLGPPAPSPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVV 240
Qy 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGFDNRFHCHFSPRDWORLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGFDNRFHCHFSPRDWORLI 300
Qy 301 NNNWGFPRKLNFKLFIQKVEVTONDGTITIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKLNFKLFIQKVEVTONDGTITIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
```

```
Qy 361 CLPPFPADVFMVPOYGYLTLLNGSQAVGRSSFYCLEYFPSPQMLRTGNNFTFSYTFEDVVF 420
Db 361 CLPPFPADVFMVPOYGYLTLLNGSQAVGRSSFYCLEYFPSPQMLRTGNNFTFSYTFEDVVF 420
Qy 421 HSSVAHSQSLDRLMNPLIDQLYLYLSRTNPSTGTTTQSRLOFSQAGASDIIDQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNPLIDQLYLYLSRTNPSTGTTTQSRLOFSQAGASDIIDQSRNWLPG 480
Qy 481 PCYRQQRVSKTSADNNNSEYSWTGTKYHLNGRDSLNVNPGPAMASHKDDKEKFFPQSGVL 540
Db 481 PCYRQQRVSKTAADNNNSEYSWTGTKYHLNGRDSLNVNPGPAMASHKDDKEKFFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNQAATAADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNQAATAADVNTQGV 600
Qy 601 LPGMVQDRDVLVLOGPIWAKIPHTDGHFHPSPLMGGFGLKHPPOQILIKNTPVPANPSTT 660
Db 601 LPGMVQDRDVLVLOGPIWAKIPHTDGHFHPSPLMGGFGLKHPPOQILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNTYKNSVNVDFTTVDNTGVY 720
Db 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNTYKNSVNVDFTTVDNTGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 8
Q6JC21
ID Q6JC21 PRELIMINARY; PRT; 735 AA.
AC Q6JC21;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530598; AAS99283.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81929 MW; 4EC10D74C451A429 CRC64;

Query Match 99.2%; Score 3961; DB 2; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.7e-240;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWMLEDTLSGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Db 1 MAADGYLPDWMLEDTLSGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Qy 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHDAEAEFOERLKEDTSFGNGLGRAVFQ 120
Db 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHDAEAEFOERLKEDTSFGNGLGRAVFQ 120
Qy 121 AKKRVLEPLGLVEEPVKTAPEGKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPEGKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAPSPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVV 240
```

```
Db 181 SVDPDPLQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Qy 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRPCHFSPRDQRLI 300
Db 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRPCHFSPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVFMVPOYGYLTNNESQAVGRSSPYCLYFPSSQMLRTGNNFTSYTFEDVVF 420
Db 361 CLPPFPADVFMVPOYGYLTNNESQAVGRSSPYCLYFPSSQMLRTGNNFTSYTFEDVVF 420
Qy 421 HSSYAHQSOLDRLMPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Db 421 HSSYAHQSOLDRLMPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Qy 541 IFGKGSSEKTNVDIEKWMITDEEBIRTTNPVATEQYGSVSTNLRGNRQAATADVNTQGV 600
Db 541 IFGKGSSEKTNVDIEKWMITDEEBIRTTNPVATEQYGSVSTNLRGNRQAATADVNTQGV 600
Qy 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMGFGFLKHPPIQILIKNTPVPANPSTT 660
Db 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMGFGFLKHPPIQILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVTDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 9
Q6JC25 PRELIMINARY; PRT; 735 AA.
AC Q6JC25, 2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Capsid protein Vp1.
GN Name=cap;
OC Adeno-associated virus.
OX Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
RN NCBI_TaxID=272636;
[1]
SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530594; AAS99279.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvocoat.
DR Pfam; PF00740; Parvocoat; 1.
DR SEQUENCE 735 AA; 81748 MW; 714DF954BF562469 CRC64;

Query Match 99.0%; Score 3955; DB 2; Length 735;
Best Local Similarity 98.9%; Pred. No. 4e-240;
Matches 727; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPKPAERHKDDSRGLVLPYKYLGPPFNGLD 60
```

```
Db 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPKPAERHKDDSRGLVLPYKYLGPPFNGLD 60
Qy 61 KGSPVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFQRLKEDTSFGGNLGRAVFQ 120
Db 61 KGSPVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFQRLKEDTSFGGNLGRAVFQ 120
Qy 121 AKKRVLLEPLGLVEEPVKTAAGKKRPVEHSPVDPDSSSGTGKAGQOPARKLNFQGTGDAD 180
Db 121 AKKRVLLEPLGLVEEPVKTAAGKKRPVEHSPVDPDSSSGTGKAGQOPARKLNFQGTGDAD 180
Qy 181 SVDPDPLQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Db 181 SVDPDPLQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Qy 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRPCHFSPRDQRLI 300
Db 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRPCHFSPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVFMVPOYGYLTNNESQAVGRSSPYCLYFPSSQMLRTGNNFTSYTFEDVVF 420
Db 361 CLPPFPADVFMVPOYGYLTNNESQAVGRSSPYCLYFPSSQMLRTGNNFTSYTFEDVVF 420
Qy 421 HSSYAHQSOLDRLMPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Db 421 HSSYAHQSOLDRLMPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Qy 541 IFGKGSSEKTNVDIEKWMITDEEBIRTTNPVATEQYGSVSTNLRGNRQAATADVNTQGV 600
Db 541 IFGKGSSEKTNVDIEKWMITDEEBIRTTNPVATEQYGSVSTNLRGNRQAATADVNTQGV 600
Qy 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMGFGFLKHPPIQILIKNTPVPANPSTT 660
Db 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMGFGFLKHPPIQILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVTDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 10
Q6JBZ5 PRELIMINARY; PRT; 735 AA.
AC Q6JBZ5,
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Capsid protein Vp1.
GN Name=cap;
OC Adeno-associated virus.
OX Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
RN NCBI_TaxID=272636;
[1]
SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530624; AAS99309.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
```

DR	GO: 0005198; F: structural molecule activity; IEA.	
DR	InterPro: IPR001403; Parvo coat.	
DR	Pfam: PF00740; Parvo coat; 1.	
SQ	SEQUENCE 735 AA; 81778 MW; 15972FCAC5764274 CRC64;	
Query Match 99.0%; Score 3954; DB 2; Length 735;		
Best Local Similarity 99.0%; Pred. No. 4.6e-240;		
Matches 728; Conservative 1; Mismatches 6; Indels 0; Gaps 0;		
Qy	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Db	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Qy	61 KGEVNEADAAAEHDKAYDRQLDSGDNPYLYKYNHADAEOFERLKEDTSFGCNLGRAVFQ 120	
Db	61 KGEVNEADAAAEHDKAYDRQLDSGDNPYLYKYNHADAEOFERLKEDTSFGCNLGRAVFQ 120	
Qy	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Db	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Qy	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGSGNWHCDSTWMDRVI 240	
Db	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGSGNWHCDSTWMDRVI 240	
Qy	241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300	
Db	241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300	
Qy	301 NNNWGFPRKLNFKLNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360	
Db	301 NNNWGFPRKLNFKLNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360	
Qy	361 CLPPFPADVFWVQGYLTLLNGSQAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420	
Db	361 CLPPFPADVFWVQGYLTLLNGSQAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420	
Qy	421 HSSVAHSQSLDRLMNPILIDQLYLSTRNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Db	421 HSSVAHSQSLDRLMNPILIDQLYLSTRNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Qy	481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQSGVL 540	
Db	481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQSGVL 540	
Qy	541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600	
Db	541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600	
Qy	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTT 660	
Db	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTT 660	
Qy	661 FSAAKFASFTQYSTGVSVSEIWEIWELOKENSKRWNPEIQTSNYSNKSNNVDFVDTNGVY 720	
Db	661 FSAAKFASFTQYSTGVSVSEIWEIWELOKENSKRWNPEIQTSNYSNKSNNVDFVDTNGVY 720	
Qy	721 SEPRPIGTRYLTRNL 735	
Db	721 SEPRPIGTRYLTRNL 735	
RESULT 11		
Q6JBZ9	PRELIMINARY; PRT; 735 AA.	
ID	Q6JBZ9	
AC	Q6JBZ9; 05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Capid protein VP1.	
GN	Namescap;	
OS	Adeno-associated virus.	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.	
OX	NCBI_TaxID=272636;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;	
RA	Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,	
RA	Wilson J.M.;	
RT	"Clades of Adeno-associated viruses are widely disseminated in human	
RT	tissues.";	
RL	J. Virol. 78:6381-6388(2004).	
DR	EMBL: AY530620; AAS99305.1: -	
DR	GO: 0019028; C: viral capsid; IEA.	
DR	GO: 0005198; F: structural molecule activity; IEA.	
DR	InterPro: IPR001403; Parvo coat.	
DR	Pfam: PF00740; Parvo coat; 1.	
SQ	SEQUENCE 735 AA; 81877 MW; 4CB8B93F79877D8C CRC64;	
Query Match 99.0%; Score 3953; DB 2; Length 735;		
Best Local Similarity 98.8%; Pred. No. 5.3e-240;		
Matches 726; Conservative 5; Mismatches 4; Indels 0; Gaps 0;		
Qy	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Db	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Qy	61 KGEVNEADAAAEHDKAYDRQLDSGDNPYLYKYNHADAEOFERLKEDTSFGCNLGRAVFQ 120	
Db	61 KGEVNEADAAAEHDKAYDRQLDSGDNPYLYKYNHADAEOFERLKEDTSFGCNLGRAVFQ 120	
Qy	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Db	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Qy	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGSGNWHCDSTWMDRVI 240	
Db	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGSGNWHCDSTWMDRVI 240	
Qy	241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300	
Db	241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300	
Qy	301 NNNWGFPRKLNFKLNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360	
Db	301 NNNWGFPRKLNFKLNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360	
Qy	361 CLPPFPADVFWVQGYLTLLNGSQAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420	
Db	361 CLPPFPADVFWVQGYLTLLNGSQAVGRSSFYCLEYFPQSOMLRTGNFTFSYTFEDVPP 420	
Qy	421 HSSVAHSQSLDRLMNPILIDQLYLSTRNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Db	421 HSSVAHSQSLDRLMNPILIDQLYLSTRNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Qy	481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQSGVL 540	
Db	481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQSGVL 540	
Qy	541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600	
Db	541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600	
Qy	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTT 660	
Db	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTT 660	
Qy	661 FSAAKFASFTQYSTGVSVSEIWEIWELOKENSKRWNPEIQTSNYSNKSNNVDFVDTNGVY 720	
Db	661 FSAAKFASFTQYSTGVSVSEIWEIWELOKENSKRWNPEIQTSNYSNKSNNVDFVDTNGVY 720	
Qy	721 SEPRPIGTRYLTRNL 735	
Db	721 SEPRPIGTRYLTRNL 735	


```
RESULT 12
Q670R4 ID Q670R4 PRELIMINARY; PRT; 735 AA.
AC Q670R4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Bartlett J.S.,
RA Sferra T.J., Shell R., Johnson P.R., Clark K.R.;
RT "Characterization of Adeno-Associated Viruses in Children.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY695373; AAU05364.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat_1.
DR SQ SEQUENCE 735 AA; 81894 MW; ECD3D9AF395F984D CRC64;

Query Match 98.8%; Score 3946; DB 2; Length 735;
Best Local Similarity 98.5%; Pred. No. 1.5e-239;
Matches 724; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGVLPGYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGVLPGYKYLGPFGNGLD 60
QY 61 KGEVPNEADAALAEHDKAYDRQLDSGDNPKYKYNHADAEPFQERLKEDTSFGNGLGRAVFQ 120
DB 61 KGEVPNEADAALAEHDKAYDRQLDSGDNPKYKYNHADAEPFQERLKEDTSFGNGLGRAVFQ 120
QY 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPSSSGTGKAGQOPARKRLNFGQTGDAD 180
DB 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPSSSGTGKAGQOPARKRLNFGQTGDAD 180
QY 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNSQNWCHDSTWMDRVI 240
DB 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNSQNWCHDSTWMDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300
DB 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
DB 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPFPADVFMVPOGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNNTFSYTFEDVFP 420
DB 361 CLPPFPADVFMVPOGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNNTFSYTFEDVFP 420
QY 421 HSSYAHQSQSLDRMLNPLIDQYLYLSTRNTPSGTTTQSRLOFSQAGASDIRQSRNLWLP 480
DB 421 HSSYAHQSQSLDRMLNPLIDQYLYLSTRNTPSGTTTQSRLOFSQAGASDIRQSRNLWLP 480
QY 481 PCVROQRVSKTSADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFPQSGVL 540
DB 481 PCVROQRVSKTSADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFPQSGVL 540
QY 541 IFKGQSEKTNVDIEKVMITDEEIRTNTPVATEQYGVSTNLQRNGRAATADVNTQGV 600
DB 541 IFKGQSEKTNVDIEKVMITDEEIRTNTPVATEQYGVSTNLQRNGRAATADVNTQGV 600
QY 601 LPMWQDDVYLGPIWAKIPIHTDGHFHPSPLMGPGFLKHPPPQILLKNTPVNPANPSTT 660
DB 601 LPMWQDDVYLGPIWAKIPIHTDGHFHPSPLMGPGFLKHPPPQILLKNTPVNPANPSTT 660
QY 661 FSAAKFASPIQTSGQSVSEIWELOKENSXRNWPEIQYTSYNKSVNVDFTVDTNGVY 720
DB 661 FSAAKFASPIQTSGQSVSEIWELOKENSXRNWPEIQYTSYNKSVNVDFTVDTNGVY 720

RESULT 13
Q6JC07 ID Q6JC07 PRELIMINARY; PRT; 735 AA.
AC Q6JC07;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RA Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530612; AAS99297.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat_1.
DR SQ SEQUENCE 735 AA; 81661 MW; 96BB878041CB47B6 CRC64;

Query Match 98.8%; Score 3945; DB 2; Length 735;
Best Local Similarity 98.8%; Pred. No. 1.7e-239;
Matches 726; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGVLPGYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGVLPGYKYLGPFGNGLD 60
QY 61 KGEVPNEADAALAEHDKAYDRQLDSGDNPKYKYNHADAEPFQERLKEDTSFGNGLGRAVFQ 120
DB 61 KGEVPNEADAALAEHDKAYDRQLDSGDNPKYKYNHADAEPFQERLKEDTSFGNGLGRAVFQ 120
QY 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPSSSGTGKAGQOPARKRLNFGQTGDAD 180
DB 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPSSSGTGKAGQOPARKRLNFGQTGDAD 180
QY 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNSQNWCHDSTWMDRVI 240
DB 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNSQNWCHDSTWMDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300
DB 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
DB 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPFPADVFMVPOGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNNTFSYTFEDVFP 420
DB 361 CLPPFPADVFMVPOGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNNTFSYTFEDVFP 420
QY 421 HSSYAHQSQSLDRMLNPLIDQYLYLSTRNTPSGTTTQSRLOFSQAGASDIRQSRNLWLP 480
DB 421 HSSYAHQSQSLDRMLNPLIDQYLYLSTRNTPSGTTTQSRLOFSQAGASDIRQSRNLWLP 480
QY 481 PCVROQRVSKTSADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFPQSGVL 540
DB 481 PCVROQRVSKTSADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFPQSGVL 540
```


Db 481 PCYRQQRVSKTADNNNSSEYWTGATKYHLNGRSLVNPGPAMASHKDDDEKFFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEYQGSVSTNLQGRNQRAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEYQGSVSTNLQSGNTQAATAADVNTQGV 600
Qy 601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPHPPQIILIKNTVPANPSTT 660
Db 601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPHPPQIILIKNTVPANPSTT 660
Qy 661 FSAAKFASFTQYSTQSGVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFTQYSTQSGVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFVTDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
RESULT 14
ID Q67006 PRELIMINARY; PRT; 735 AA.
AC Q67006;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
SEQUENCE FROM N.A.
RA Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Bartlett J.S.,
RA Sferri T.J., Shell R., Johnson P.R., Clark K.R.;
RT "Characterization of Adeno-Associated Viruses In Children."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY695378; AAU05372.1; -.
DR EMBL; AY695374; AAU05366.1; -.
DR InterPro; IPR001403; Parvo coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00740; Parvo_coat_1.
SQ SEQUENCE 735 AA; 81784 MW; 818BBE497DA91984 CRC64;
Query Match 98.7%; Score 3941; DB 2; Length 735;
Best Local Similarity 98.4%; Pred. No. 3e-239;
Matches 723; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGEPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Db 61 KGEPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSPVEPDSSTGKAGQOQPAKRLNFQGTGDAD 180
Db 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSPVEPDSSTGKAGQOQPAKRLNFQGTGDAD 180
Qy 181 SVDPDPQLGPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
Db 181 SVDPDPQLGPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
Qy 241 TTSTRWALPTYNHLYKQISSQSGASNDHYGYSTPWGYFDENRPHCFSPRDWQRLI 300
Db 241 TTSTRWALPTYNHLYKQISSQSGASNDHYGYSTPWGYFDENRPHCFSPRDWQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANLNTSTVQVFTDSYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANLNTSTVQVFTDSYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVPVMPQGYLTLNNGSQAVGRSSFYCLEYFPQSMLRTGNFTSFYEDVFP 420

Db 361 CLPPFPADVPVMPQGYLTLNNGSQAVGRSSFYCLEYFPQSMLRTGNFTSFYEDVFP 420
Qy 421 HSSVAHSQSIDRLMNPILIDVLYLSRTNTPSGTTTOSRLQFSQAGASDIRDQRNWLPG 480
Db 421 HSSVAHSQSIDRLMNPILIDVLYLSRTNTPSGTTTMSRLQFSQAGASDIRDQRNWLPG 480
Qy 481 PCYRQQRVSKTADNNNSSEYWTGATKYHLNGRSLVNPGPAMASHKDDDEKFFPQSGVL 540
Db 481 PCYRQQRVSKTADNNNSSEYWTGATKYHLNGRSLVNPGPAMASHKDDDEKFFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEYQGSVSTNLQGRNQRAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEYQGSVSTNLQSGNTQAATAADVNTQGV 600
Qy 601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPHPPQIILIKNTVPANPSTT 660
Db 601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPHPPQIILIKNTVPANPSTT 660
Qy 661 FSAAKFASFTQYSTQSGVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFTQYSTQSGVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFVTDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
RESULT 15
ID Q6JC05 PRELIMINARY; PRT; 735 AA.
AC Q6JC05;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilison J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530614; AAS99299.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat_1.
SQ SEQUENCE 735 AA; 81755 MW; E1588D5235171373 CRC64;
Query Match 98.6%; Score 3940; DB 2; Length 735;
Best Local Similarity 98.9%; Pred. No. 3.5e-239;
Matches 727; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGEPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Db 61 KGEPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSPVEPDSSTGKAGQOQPAKRLNFQGTGDAD 180
Db 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSPVEPDSSTGKAGQOQPAKRLNFQGTGDAD 180
Qy 181 SVDPDPQLGPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
Db 181 SVDPDPQLGPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240

```

Db 181 SVPDPQLGQPPAAPSGLGTNTMATGSGAPWADNNEGADGVGSSGNRHCDSITWMDRVI 240
Qy 241 TTSTRTWALPTYNHLYKQISSQSGANDNHYFGYSTPWGYFDNRFHCHFSPRDWORLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGANDNHYFGYSTPWGYFDNRFHCHFSPRDWORLI 300
Qy 301 NNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVMPVQGYHITLNGSQAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEDVPF 420
Db 361 CLPPFPADVMPVQGYHITLNGSQAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEDVPF 420
Qy 421 HSSYAHQSGLDRLANPLIDQYLYLSRTNTPSGTTQSRLOFSQAGASDIRDOSRNWLP 480
Db 421 HSSYAHQSGLDRLANPLIDQYLYLSRTNTPSGTTQSRLOFSQAGASDIRDOSRNWLP 480
Qy 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEKFFPOSGLV 540
Db 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEKFFPOSGLV 540
Qy 541 IFGKQSEKTNVDIEKVMITDEEIRTNVPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Db 541 IFGKQSEKTNVDIEKVMITDEEIRTNVPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Qy 601 LPGMVQDRDQVYLOQPIWAKIPTHGDFHPSPLMGFGPLKHPPOILLIKNTVPANPSTT 660
Db 601 LPGMVQDRDQVYLOQPIWAKIPTHGDFHPSPLMGFGPLKHPPOILLIKNTVPANPSTT 660
Qy 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

```

Search completed: May 19, 2005, 13:32:59
 Job time : 175 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:16:50 ; Search time 44 Seconds
(without alignments)
1607.257 Million cell updates/sec

Title: US-10-038-972A-13
Perfect score: 3994
Sequence: 1 MAADGYLPDWLEDTLSEGR.....TNGVYSEPRPGTRYLFRNL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2548	63.8	504	1 VCPV3A	coat protein - ade
2	2122	53.1	732	2 S52210	coat protein VP1 -
3	781.5	19.6	673	1 VCPVB5	coat protein VP1 -
4	602	15.1	781	1 VCPV19	coat protein VP1 -
5	504.5	12.6	729	1 A60006	coat protein VP1 -
6	501.5	12.6	729	1 VCPVNA	coat protein VP1 -
7	461.5	11.6	727	1 VCPV1F	coat protein VP1 -
8	457.5	11.5	727	1 VCPVFP	coat protein VP1 -
9	452.5	11.3	722	1 VCPVME	coat protein VP1 -
10	442	11.1	718	1 VCPVIM	coat protein VP1 -
11	442	11.1	748	1 VCPVCP	coat protein VP1 -
12	434.5	10.9	723	1 VCPVPP	coat protein VP1 -
13	425.5	10.7	737	1 VCPVCD	coat protein VP1 -
14	411	10.3	722	1 VCPVCN	coat protein VP1 -
15	397	9.9	722	1 VCPV2	coat protein VP1 -
16	386.5	9.7	716	1 VCPV2M	coat protein VP1 -
17	232.5	5.8	587	1 B44276	coat protein VP1 -
18	213	5.3	584	2 S49594	capsid protein VP2
19	194	4.9	810	2 A44054	orf1 protein - Jun
20	170	4.3	702	1 VCPVAP	coat protein VP1 -
21	133	3.3	690	2 AB0124	probable TonB-depe
22	132.5	3.3	1129	2 T25635	hypothetical prote
23	129	3.2	743	2 T09173	EH domain protein
24	128	3.2	2529	2 B64635	toxin-like outer m
25	124.5	3.1	1145	2 T18235	transcription acti
26	122.5	3.1	1175	2 T25634	hypothetical prote
27	122	3.1	2399	2 H71879	toxin-like outer m
28	121.5	3.0	1745	2 A46431	tight junction-ass
29	121.5	3.0	2500	2 G71609	hypothetical prote

30 119.5 3.0 803 2 A86655
31 119.5 3.0 1213 2 S16356
32 119 3.0 932 2 S62555
33 119 3.0 1742 2 S76110
34 118.5 3.0 1159 2 S62562
35 118 3.0 1858 2 T18273
36 117.5 2.9 2894 2 C64474
37 117 2.9 833 2 A47528
38 117 2.9 866 2 D44234
39 117 2.9 1344 1 A35175
40 115.5 2.9 309 2 I48125
41 115.5 2.9 925 2 T19361
42 115 2.9 931 2 T49710
43 115 2.9 1049 2 A27079
44 114.5 2.9 894 2 S51245
45 114 2.9 1403 2 T11583

hypothetical prote
ovo protein - frui
protoplast regene
hypothetical prote
probable nuclear p
i-phosphatidylinos
hypothetical prote
transcription fact
fibrinogen alpha c
mucin 1 precursor,
syndecan-1 - Chine
hypothetical prote
related to glucan
fibronectin recept
probable finger pr
probable translati

coat protein - adeno-associated virus type 2
C;Species: adeno-associated virus type 2
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C;Accession: A03698
R;Strivastava, A.; Lusby, E.W.; Berns, K.I.
J. Virol. 45, 555-564, 1983
A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.
A;Reference number: A03694; MUID:83164299; PMID:6300419
A;Accession: A03698
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-504 <SRI>
A;Cross-references: UNIPROT:P03135; EMBL:J01901; NID:G209616; PIDN:AAA42376.1; PID:G209616
C;Superfamily: adeno-associated virus coat protein
C;Keywords: coat protein

ALIGNMENTS

RESULT 1

VCPV3A

coat protein - adeno-associated virus type 2

C;Species: adeno-associated virus type 2

C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004

C;Accession: A03698

R;Strivastava, A.; Lusby, E.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A;Reference number: A03694; MUID:83164299; PMID:6300419

A;Accession: A03698

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-504 <SRI>

A;Cross-references: UNIPROT:P03135; EMBL:J01901; NID:G209616; PIDN:AAA42376.1; PID:G209616

C;Superfamily: adeno-associated virus coat protein

C;Keywords: coat protein

Query Match 63.8%; Score 2548; DB 1; Length 504;

Best Local Similarity 99.2%; Pred. No. 3.9e-159;

Matches 470; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Qy	203	MATGSGAPMADNNEGADGVGNSGNHCHDSTWMDGRVITTTSTRTWALPTYNHLYKQISS	262
Db	1	MATGSGAPMADNNEGADGVGNSGNHCHDSTWMDGRVITTTSTRTWALPTYNHLYKQISS	60
Qy	263	QSGASNDNHYPGYSTPWGYFDNRFCHFSPRDQORLIINNNGWFRPKRLNFIOVKE	322
Db	61	QSGASNDNHYPGYSTPWGYFDNRFCHFSPRDQORLIINNNGWFRPKRLNFIOVKE	120
Qy	323	VTQNDGTTTIANNTLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVPMVPOYGLTLNN	382
Db	121	VTQNDGTTTIANNTLTSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVPMVPOYGLTLNN	180
Qy	383	GSQAVGRSSFFCYLBYFSPQMLRTGNFTFSYTFEDVPHSSYAHSSQSLDRMLNPLIDQYL	442
Db	181	GSQAVGRSSFFCYLBYFSPQMLRTGNFTFSYTFEDVPHSSYAHSSQSLDRMLNPLIDQYL	240
Qy	443	YYLRTNTPSGTTTQSRLOFSQAGASIRDOSRNWLPGPCYRQORVSKTSADNNNSYSW	502
Db	241	YYLRTNTPSGTTTQSRLOFSQAGASIRDOSRNWLPGPCYRQORVSKTSADNNNSYSW	300
Qy	503	TGATKYHLNGRDSLVPNPGPAMASHKDEKFPFQSGVLIIFGQSGSEKTNVDIEKVMITDE	562
Db	301	TGATKYHLNGRDSLVPNPGPAMASHKDEKFPFQSGVLIIFGQSGSEKTNVDIEKVMITDE	358
Qy	563	EEIRTTNPVATEQYGSYSTNLQRNQAAATADVNTQVLPQMWQDRDVIYLGPIWAKIP	622
Db	359	EEIGTTNPVATEQYGSYSTNLQRNQAAATADVNTQVLPQMWQDRDVIYLGPIWAKIP	418
Qy	623	HTDGHFSPMLMGFGGLKHPPEQLIKNTVPANPSTTFFSAKFPASFTQYSTG	676

Db 419 HTDGHFHPPLMGFGGLKHPHPQILLIKNTVPANPSTTFSAAKFPASFTIQYSTG 472

RESULT 2
S52210
N:Alternate names: VP1 protein
C:Species: muscovy duck parvovirus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C/Accession: S52210
R/Zadori, Z.; Erdel, J.; Nagy, J.; Kisari, J.
submitted to the EMBL Data Library, September 1993
A:Reference number: S52209
A/Accession: S52210
A/Molecule type: DNA
A/Residues: 1-732 <2AD>
A/Cross-references: UNIPROT:Q83289; EMBL:X75093; NID:G609091; PIDN:CAAS2984.1; PID:G609090
A/Experimental source: strain FM
C:Genetics:
A:Gene: VP1
C:Superfamily: parvovirus coat protein
C/Keywords: coat protein

Query Match 53.1%; Score 2122; DB 2; Length 732;
Best Local Similarity 54.2%; Pred. No. 5.2e-131;
Matches 409; Conservative 92; Mismatches 194; Indels 60; Gaps 12;

QY 9 DWLEDTLSEGIQWKLKPGPPPKPAE-----RHKDSRGLVLPFGYKYLGFEN 57
Db 10 DWYETAAA-----SMRHLKAGAPKPKSNQOSQSVSTRDKPQRKNNRGVLPFGYKYLGFEN 65

QY 58 GLDKGEPVNEADAALAEHDKAYDRLDSDGNPNLYKYNHADAEOERLKEDTSFGNIGRA 117
Db 66 GLDKGPPVNAKADSVALEHDKAYDQQLKAGDNPYIKFKAHQDEFIDNLQDTSFGNIGRA 125

QY 118 VFQAKRVLEPLGLVEEVPKTAPOKK-----RPVESHV-----EPDSSSGTGKAGQOQA 167
Db 126 VFQAKRVLEPLGLVEEVPKTAPOKK-----RPVESHV-----EPDSSSGTGKAGQOQA 167

QY 168 RKRLNFGTGADSVDPDQPLGQPPAAPSGLGTTNTMATGSGAPMADNNEGADGVGNSSGN 227
Db 186 AATSG-----SEPVAPN-----MAEGSGMGDSAGGADGVGNASGN 223

QY 228 WHCDSTWMDRVITSTRWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPFNRF 287
Db 224 WHCDSTWMDRVITSTRWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPFNRF 287

QY 288 HCHFSRDMQRLINNNWGFPRKLNFKLFWIOUKEVTQNDGTTIANLSTVQVTFDSE 347
Db 284 HCHFSRDMQRLINNNWGFPRKLNFKLFWIOUKEVTQNDGTTIANLSTVQVTFDSE 347

QY 348 YQLPYVLSAQHGGCLPPFPADVFMVPOYGYLTN---NGSQAVGRSFFCYCLEYFPQMLR 404
Db 344 HQLPYVLSAQHGGCLPPFPADVFMVPOYGYLTN---NGSQAVGRSFFCYCLEYFPQMLR 404

QY 405 TGNFTSYTFEDVPFHSSVAHSQSLDRLNPLIDQVLYLSTNTTPSGTTQSRLOFSQ 464
Db 404 TGNFTSYTFEDVPFHSSVAHSQSLDRLNPLIDQVLYLSTNTTPSGTTQSRLOFSQ 464

QY 465 AGASDIRDQSRNMLPGPCYQQRVSKTSADNNN-SEYS-WTGATKYHLNGRSLVNPQA 522
Db 458 AVKGFAGMGRNMLPGKLLDQVRAYSQGTGTDNYANWSIWSKNVFLKDKREYLLQPGPV 517

QY 523 MASHKODEEFPQSGVLIQKQ--GSEKTNVDIEKVMITDEBEIRTNVPATEQYGSVS 580
Db 518 ATTHTEDQASSVPAQNIIGAKDPYRSGSTLAGISDMVTDEQEIAPTNGVGNRPYGLTV 577

QY 581 TNLQGRNRQAATADVNTQGVLPQWMDRDVYLGPIWAKIPHTDGHFHPPLMGFGGLK 640
Db 578 TNEQNTTAPTNAELVGLGALPGWQNRDIYQGGPIWAKIPKTDGKPHPSNLOGFGLH 637

QY 641 HPPQILLIKNTVPANPSTTFSAAKFPASFTIQYSTQSVSEIWELOKENSKRWNPEIQY 700
Db 641 HPPQILLIKNTVPANPSTTFSAAKFPASFTIQYSTQSVSEIWELOKENSKRWNPEIQY 700

Db 638 NPPQVFEIKNTVPADPPLEVNVQKWSYITQYSTGQCTVMYMWELKENSKRWNPEIQF 697

QY 701 TSNYKNSVNDFTVDYNGVYSEPRPIGTTRYLTNLL 735
Db 698 TSNFNGRTSTWFPAPNETGGYVEDRLIGTRYLTQNL 732

RESULT 3
VCPVB5
N:Contains: VP1 - bovine parvovirus
C:Species: bovine parvovirus
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: A26104
R/Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.
J. Virol. 60, 1085-1097, 1986
A:Title: Complete nucleotide sequence and genome organization of bovine parvovirus.
A/Reference number: A26104; MUID:87061184; PMID:3783814
A/Accession: A26104
A/Molecule type: DNA
A/Residues: 1-673 <CHE>
A/Cross-references: UNIPROT:P07297; EMBL:M14363; NID:G333454; PIDN:AAB59847.1; PID:G80888
C:Superfamily: parvovirus coat protein
C/Keywords: coat protein
F/138-673/Product: coat protein VP2 #status Predicted <VP2>

Query Match 19.6%; Score 781.5; DB 1; Length 673;
Best Local Similarity 29.2%; Pred. No. 2.6e-43;
Matches 213; Conservative 104; Mismatches 299; Indels 113; Gaps 22;

QY 30 PPPKPAERHKDSRGLVLPFGYKYLGFENADAAALAEHDKAYDRLDSDGNP 89
Db 2 PPTNKA-----NSKKGLTLPFGYKYLGFENADAAALAEHDKAYDRLDSDGNP 89

QY 90 YLKYNHADAEOERLKEDTSFGNIGRAVFOAKRVLEPLGLVEEVPKTAPOKKPVEHS 149
Db 58 YLYFNTHDQNLDELKDDTSFGKLGARGVQIKKALAPAL-----POTSGKGRDALRRK 111

QY 150 PVEPDSSSGTGKAGQOQAPARKLNFGQTDGADSVDPDQPLGQPPAAPSGLGTTNTMATGSGA 209
Db 112 LYFARSNKAKKANREPAPSTSNQNMVEVNDIPNDEAGNQ---PIELATRSV-VGSGS 166

QY 210 PMADNNEGADGVGNSSGNWCHDSTWMDRVITSTRWALPTYNHLYKQISSQSGASND 269
Db 167 -VGCGRGGSGYGTGGTGTTFSENIVTKTRQICDKNGHLYKSEVLNTGDTAH 225

QY 270 NHFGYSTPWGYPFNRFHCHFSRDMQRLINNNWGFPRKLNFKLFWIOUKEVTQNDGT 329
Db 226 RQY-AITTPWSYFNPNQYSSHFSPNDWQHLVNDYERFRPKAMIVRVYNLQIKQIMTDGAM 284

QY 330 TTIANN-LTSTVQVTFDSEYQLPYVLSAQHGGCLPPFPADVFMVPOYGYLTN-----NN 382
Db 285 GTVYNNDLTAGMHIFCDGDHRYPYVQHWPDDQCPNLSIWELPQYAYIPAPISVVDDN 344

QY 383 GSQAVGRS-----SPYCLEYFPQMLRTGNFTSYTFEDVPFHSSVAHSQ--SLDRLMNP 436
Db 345 TTNVTEHLLKGVPLMYLMSDHEVLNNGRIYRIYIQLWRLMDRKQHHIQHASDDVQST 404

QY 437 LIDQVLYLSTNTTPSGTTQSRLOFSQAGASDIRDQSRNMLPGPCYQQRVSKTSADNN 496
Db 405 GQKQKLLIORTKQP-----NKORFONAAAL-----RTSNMWSGF-----GIARGTHNAT 448

QY 497 NSEYSWTGATKYHLNGRSLVNPQPMASHKDDBEKFFQSGVL-----IFGKQG 546
Db 449 LQTSAGALVTMTNGAD-----VSGVRAVRGVSTDPDIYGGQQ 487

QY 547 SEKTNVDIEKVMITDEBEIRTNVPATEQYGSVSTNLQRG--NRQAAT-----AD 594
Db 488 PE-----SDLLRLRYASAAGQQNPILNAAARHTFTREARTKLTGNSGADGD 536

QY 595 VNTQGVLPQWMDRDVYLGPIWAKIPHTDGHFHPPLMGFGGLKHPHPQILLIK--NTP 652
Db 537 YKEMWMLPNQWDSAPISRSNPINVKRVNRNKTLLDTQDGSIPMSHPGTGTFIKLARIP 596


```

Db      545  TFFNQSKGRAPKQFNOQAPLNLENTNNGTLLPSDPIGKPNMHPMNTLNTYGLTALN 604
Qy      606  WQDRDVLQGIWAKIPIHTD---GHFHPSPLMGGFGLK-HPPQILIKNTVPANPSTFS 662
Db      605  -NTAPVPNGQIWDKELDTLKLPLH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFN 657
Qy      663  A-AKFASFIQYSTGQSVSEIWELOKENSKRNPPEIQYTSNYSNKSNNVDFVDTN--GV 719
Db      658  ADSFQPRIIITYSNFWKGTLTFTAKMRSSNMNPIQOHTTT---AENIGNIYPTNIGGI 714
Qy      720  -----YSE--PRPI 726
Db      715  KMPPEYSQLIPRKL 728

RESULT 6
VCPVNA
N;Contains: coat protein VP2
C;Species: porcine parvovirus
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: B33743; D48472
R;Vasudevacharya, J.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 173, 368-377, 1989
A;Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pa
A;Reference number: A33743; MUID:90085785; PMID:2596019
A;Accession: B33743
A;Molecule type: DNA
A;Residues: 1-729 <VAS>
A;Cross-references: UNIPROT:P18546; GB:M32787; NID:G332983; PIDN:AAA46917.1; PID:G332985
R;Bergerson, J.; Menezes, J.; Tijssen, P.
Virology 197, 86-98, 1993
A;Title: Genomic organization and mapping of transcription and translation products of p
A;Reference number: A48472; MUID:94025614; PMID:8212598
A;Accession: D48472
A;Molecule type: DNA
A;Residues: 11-729 <BER>
A;Experimental source: NADL-2, ATCC VR-742
A;Note: sequence extracted from NCBI backbone (NCBIN:138789, NCBIPI:138794)
C;Genetics: 10/1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein; glycoprotein
F;151-729/Product: coat protein VP2 #status predicted <VP2>
F;32,172,198,282,330,433,471,573,604,651/Binding site: carbohydrate (Aen) (covalent) #eq

Query Match 12.6%; Score 501.5; DB 1; Length 729;
Best Local Similarity 26.4%; Pred. No. 6.1e-25;
Matches 209; Conservative 103; Mismatches 318; Indels 163; Gaps 36;

Qy      30  PPPKPAERHKDDSRGLVLPQYKYLGPNGLDKGEFVNEADAALHDKAYDRQLDSDNP 89
Db      3  PPAKRA-----RGLTLPQYKYLGPNSLDQGEPTNPSSDAAAKEHDEAYDKYIKGKNP 55
Qy      90  YLYNHADAEFQRLKEDTSFGGNLGRAVFOAKRVLEPLGLVVEPVKTAPEGKRPVPHS 149
Db      56  YFYFSADEAFIKETEHAKDYGGKIGHYFRAKRAFAKLSETDSTPTS---QOPEVRRS 112
Qy      150  PVPEPDSSSGTGK-----AGQOPARKRLNFGTGDADSDVPDPQPLGQPPAAPSGLG 199
Db      113  PRKHGSKPPKRPAPRPHIFINLAKKAKGTSNTNSMSSENVQHNPI-----NAG 164
Qy      200  TMTMTGTS---GAPMADNNEGADGVGNSSG--NWHCDSTWMGD---RVITTSRTWALPT 251
Db      165  TELSATNESGGGGGGGKAGCGVSTGTFNNQTFYLGGLGVRIITAHARLIHLNM 224
Qy      252  YNNHLYKQI---SSQSCAS----NDNHYFGYSTPWGDFNFRPHCFSPRDWQRLINNWN 304
Db      225  PEHETVKRIHVLNSESQVACQVQDDAHTQMVTFWSLIDANAGVWFNPADWQLISNNWT 284
Qy      305  GPRPKRLNFKLNIQKEVTQN--DGTITIANN-LISTVQVFTDSEYQLPYVLGSAHQGC 361

```

```

Db      285  EINLVSEFEQEIFPNVLKTTITESATSPPTKIYNDLTASLMVALDNTNTLPTTPAARSET 344
Qy      362  LPFPFADVFMVPOYGYL-----TLNNGSQAVGRS-----SFYCYCLE-YPPSQ 401
Db      345  LGFYPLPTKPTQYRYVLSCLRNLPPTVTGQSQQTDSIQTGLHSDIMFYTIENAVPIH 404
Qy      402  MLRTGNNE-TFSTYTEDVPFHSYAHQSQSLDR-----LNNPLI--DOYLYLSTRTNT 450
Db      405  LLRTGDEFSTGIYHFTKPL--KLTHSWQTNRSLSGLPPKLLTEPTTTEGQHQHPTLPAANT 462
Qy      451  PSCTTTTQSRLLQFQAGASDIRQSRNWLPGPCYRQORVSKTSADNNNSSEYMTGATKYHL 510
Db      463  RKGYHQTINNSYTEATA--IRPAQVGY-----NTPYNNPEYS-----497
Qy      511  NGRDSLNVNPGPAMASHKDEEKFPPQSGVLIIFGQKSEKTNVDIEKVMIT---DEBEIRT 567
Db      498  NGGPFLLTPIVPTADTQYNDDE-----PNGAIRFTMDYQGHLLTTSQELERYT 545
Qy      568  TNP-----VATROYG-SVSTNLQRGNROAATD-----VNTQGVLPGMVW 606
Db      546  FNPQSKGRAPKQFNOQAPLNLENTNNGTLLPSDPIGKSNMHPMNTLNTYGLTALN- 604
Qy      607  QDRDVLQGIWAKIPIHTD---GHFHPSPLMGGFGLK-HPPQILIKNTVPANPSTTFS 663
Db      605  NTAPVPNGQIWDKELDTLKLPLH---VTAPFVCKNPPGQLFVKIAP---NLTDDEFN 658
Qy      664  -AKFASFIQYSTGQSVSEIWELOKENSKRNPPEIQYTSNYSNKSNNVDFVDTN--GV- 719
Db      659  DSFQPRIIITYSNFWKGTLTFTAKMRSSNMNPIQOHTTT---AENIGNIYPTNIGGIR 715
Qy      720  -----YSE--PRPI 726
Db      716  MPPEYSQLIPRKL 728

RESULT 7
VCPVIF
N;Contains: coat protein VP1 - feline panleukopenia virus
C;Species: feline panleukopenia virus, FPLV
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C;Accession: A03701
R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parv
A;Reference number: A03697; MUID:85265017; PMID:2991581
A;Accession: A03701
A;Molecule type: DNA
A;Residues: 1-727 <CAR>
A;Cross-references: UNIPROT:P04864; EMBL:M10824; NID:G333474; PIDN:AAA47161.1; PID:G33347
C;Genetics: 11/1
A;Introns: 11/1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.6%; Score 461.5; DB 1; Length 727;
Best Local Similarity 24.6%; Pred. No. 2.5e-22;
Matches 191; Conservative 102; Mismatches 310; Indels 173; Gaps 32;

Qy      30  PPPKPAERHKDDSRGLVLPQYKYLGPNGLDKGEFVNEADAALHDKAYDRQLDSDNP 89
Db      3  PPAKRA-----RGLVPPGYKYLGPNSLDQGEPTNPSSDAAAKEHDEAYAYLRSGKNP 56
Qy      90  YLYNHADAEFQRLKEDTSFGGNLGRAVFOAKRVLEPL-GLVEEPVKTAPGKRPVPH 148
Db      57  YLYFSPADQRFIDQTKDATDWGKGKIGHYFRAKKAIAPVLTDPDHPSTSRPTK--PTRK 114
Qy      149  SPVEPDSSSGTGKAGQOPARKRLNFGTGDADSDVP-----DPOPLGQPPAAPSGLGTNTWA 204
Db      115  SKPPPHIFINLAK-----KKKAGAGQVKRDNQAPMSDGAQVDDGQPAV-----RNERA 163
Qy      205  TGSG-APMADNNEGADGVGNSSGNWH-----CDSTWMGDRTVITTTTRTWTALPTNNHL 256

```

Db 164 TGSNGSGGGGGGGVIGISTGFNNQTEFKFLENGWV--EITANSRLVHLNMPESN 221
Qy 257 YKQI-----SSQSGASNDNHFGYSTPGWYDFDNRFHCFSPDRDQRLNNWGRP 308
Db 222 YKRVVNNMDKTAVKGNMALLDTHVQIVTPNSLVDANAWGVWFPGDWQLIVNTMSELH 281
Qy 309 KRLNFKLNIQVKEVTQN---DGTITIANNTSTVQVTDSEYQLPYVLGSAHQCL--- 362
Db 282 VSFQEIFNVVLKTVSESATQPTKVNNDLTASLWALDSNNTMPTPAARSETLGFY 341
Qy 363 ---PPPPA-----DVFMVPOQYGLTLNNGSOAVGRS-----SFYCLE-YPPS 400
Db 342 PWKPTIPTPMRYFFQWDRTLIPSH-----TGTSGTPTNIYHGTDPPDDVQFYTIENSVPV 395
Qy 401 QMLRTGNF--TFSYTFEDVPFHSSYAHQSQSLDRLMN--PLIDQVLYLSRSTNTPS-GTTT 456
Db 396 HLLATGDEFATGTFEFDCKP--CRLTHWTQNRALGLPFLNSLPQSEGATNFGDIGVQ 453
Qy 457 QSRLOFSQAGASDIRDOSRNWLPQPCVQRQVRKTSADNNNSEYSWTGATKYHLNGRDSL 516
Db 454 DKRRGVITQMGNTDYTEATIMRP-----AEGYSAPYYSFEASTQ----- 493
Qy 517 VNPGR-----AMASHKDDREKPPQSGVLIFFGQKSEKTNVDIE----- 555
Db 494 ---GPFKIPIAAGRGGAQTDENQAADGDPFYAFGRHQGKTTTGTETPERFTYIAHQDTG 550
Qy 556 -----KVMITDEEIRTNPNVATEQVGSVSTNLQRNQAATADVNTQGVLP 602
Db 551 RYPAGDWIQINFLPVTNDVLLPTDPIG-----GKTGINY-----TNIFTYIGPLT 598
Qy 603 GMMVQDRDVLQGPWIAKIPHTDGHFSPMLMGFGGLKHP-----PPQILLIKNTVPVA 655
Db 599 ALN-NVPPVYPNGQIWDKEPDTD-----LKPRLHVNAFPVCQNCQGLFVKVAPNLT 650
Qy 656 NPSTTFAAKFASFIQYSTQGVSVIEWELQENSKRWNPETQYTSNYSKSNVND 711
Db 651 NEYDPPDASANMSRIVT-YSDFWFKGKLVFKAKLRASHTWNP-IQ-----QMSINVD 699

RESULT 8
VCPVFP
N;Contains: coat protein VP2
C;Species: feline panleukopenia virus, FPLV
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: B36608
R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus
A;Reference number: A36608; MUID:91073139; PMID:2174965
A;Accession: B36608
A;Molecule type: DNA
A;Residues: 1-727 <MAR>
A;Cross-references: UNIPROT:P24840; GB:X55115
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.5%; Score 457.5; DB 1; Length 727;
Best Local Similarity 24.8%; Pred. No. 4.6e-22;
Matches 193; Conservative 99; Mismatches 310; Indels 175; Gaps 33;

Qy 30 PPPPAERHKDDSLGLVLPGYKYLGPNGLDKGFVNEADAAALHDKAYDRQLDSDGNP 89
Db 3 PPAKRR-----RGLVPPGYKYLGPNGSLDQGEFTNPSDAAAKEHDEAYAAVLRSGNP 56

Qy 90 YLKNHADAERFQERLKEDTSFGGNLGRAVFOAKKRVLEPL-GLVEEPPVKTAGPKRVPVEH 148
Db 57 YLYFSPADQRFIDQTKADKDWGGKIGHYFFRAKKAIAPLVTDTPDHSTSRPTK--PTKR 114

Qy 149 SPVEPDSSTGK-----AGQPARKRLNFQGTGDADSVDPQPLGQPPAAPSGLGNTM 203

Db 115 SKPPPHPIINLAKKKKAGAGQV---KRDNLAPMSDGAQVQPDG---GQPAV-----RNER 162
Qy 204 ATGSG-APMADNNEGADGVGNSGNWH-----CDSTMGDRVITVSTRTWALPTYNH 255
Db 163 ATGSGNGSGGGGGGGVIGISTGFNNQTEFKFLENGWV--EITANSRLVHLNMPSE 220
Qy 256 LYKQI-----SSQSGASNDNHFGYSTPGWYDFDNRFHCFSPDRDQRLNNWGRP 307
Db 221 NYKRVVNNMDKTAVKGNMALLDTHVQIVTPNSLVDANAWGVWFPGDWQLIVNTMSELH 280
Qy 308 PKRLNFKLNIQVKEVTQN---DGTITIANNTSTVQVTDSEYQLPYVLGSAHQCL--- 362
Db 281 LVSEFQEIFNVVLKTVSESATQPTKVNNDLTASLWALDSNNTMPTPAARSETLGF 340
Qy 363 ---PPPPA-----DVFMVPOQYGLTLNNGSOAVGRS-----SFYCLE-YFP 399
Db 341 YPWKPTIPTPMRYFFQWDRTLIPSH-----TGTSGTPTNIYHGTDPPDDVQFYTIENSVP 394
Qy 400 QMLRTGNF--TFSYTFEDVPFHSSYAHQSQSLDRLMN--PLIDQVLYLSRSTNTPS-GTT 455
Db 395 VHLRTGDEFATGTFEFDCKP--CRLTHWTQNRALGLPFLNSLPQSEGATNFGDIGVQ 452
Qy 456 TQSRLOFSQAGASDIRDOSRNWLPQPCVQRQVRKTSADNNNSEYSWTGATKYHLNGRDS 515
Db 453 QDKRRGVITQMGNTDYTEATIMRP-----AEGYSAPYYSFEASTQ----- 493
Qy 516 LVNPGP-----AMASHKDDREKPPQSGVLIFFGQKSEKTNVDIE----- 555
Db 494 ---GPFKIPIAAGRGGAQTDENQAADGDPFYAFGRHQGKTTTGTETPERFTYIAHQDT 549
Qy 556 -----KVMITDEEIRTNPNVATEQVGSVSTNLQRNQAATADVNTQGVLP 601
Db 550 GRYPEGDWIQINFLPVTNDVLLPTDPIG-----GKTGINY-----TNIFTYIGPL 597
Qy 602 GMMVQDRDVLQGPWIAKIPHTDGHFSPMLMGFGGLKHP-----PPQILLIKNTVPV 654
Db 598 TALN-NVPPVYPNGQIWDKEPDTD-----LKPRLHVNAFPVCQNCQGLFVKVAPNLT 649
Qy 655 ANPSTTFAAKFASFIQYSTQGVSVIEWELQENSKRWNPETQYTSNYSKSNVND 711
Db 650 TNEYDPPDASANMSRIVT-YSDFWFKGKLVFKAKLRASHTWNP-IQ-----QMSINVD 699

RESULT 9
VCPVME
N;Contains: coat protein VP1 - mink enteritis virus (strain Abashiri)
C;Species: mink enteritis virus, MEV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: B38350
R;Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinag
J. Gen. Virol. 72, 867-875, 1991
A;Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A;Reference number: A38350; MUID:91202123; PMID:2016597
A;Accession: B38350
A;Molecule type: DNA
A;Residues: 1-722 <KAR>
A;Cross-references: UNIPROT:P27437; GB:D00765
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;139-722/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.3%; Score 452.5; DB 1; Length 722;
Best Local Similarity 24.8%; Pred. No. 9.7e-22;
Matches 189; Conservative 99; Mismatches 306; Indels 169; Gaps 32;

Qy 44 GLVLPGYKYLGPNGLDKGFVNEADAAALHDKAYDRQLDSDGNPYLKNHADAERFQER 103
Db 6 GLVPPGYKYLGPNGSLDQGEFTNPSDAAAKEHDEAYAAVLRSGKNPYLYFSPADQRFIDQ 65

Qy 104 LKEDTSFGGNLGRAVFOAKKRVLEPL-GLVEEPPVKTAGPKRVPVEHSPVDSSTGK- 161
Db 66 TKOATDWGGKIGHYFFRAKKAIAPLVTDTPDNPSRPTK--PTKRSKPPPHIFINLAKK 123


```

QY 162 ----AGQPARKELNFGTGDADSDVDPDQPPQLGPPAAPSGGLTNTMATGSG-APMADNNE 216
Db 124 KKAGAGQV-----KRDNLAPMSDGAQVDPG-----GQPAV-----RNERATGSGSGGGGG 171
QY 217 GADGVGNSGNWH-----CDSWMDRVITTSRTWALTNYNNHLYKQI-----S 261
Db 172 GSGGVGISTGTNNQTEPKFLENGWV--EITANGSRLVHLNMPSENYKRVVNNMDKTA 229
QY 262 SOSGASNDNHFGYSTPGYDFDNRHCHFSRDMQRLINNMMFRPKRLNFKLENIOVK 321
Db 230 VKGNMALDTHVQIVTPMNSLVDAWANGWFNPGDWQLIVNTMSLHVSFEQIFNVULK 289
QY 322 EYTON---DGTITANNITSTVQVFTDSEYQLPVLSAHOGL-----PPPPA----- 367
Db 290 TVSESATQPTKVYNNDLTASLWALDSNNTPTPAAMRSETLGFYPMKPTIPTPWRY 349
QY 368 ---DVFVMPQGYLTNNGSQAQVRS-----SFYCLE-YFPSSQMLRTGNPF-TFS 412
Db 350 FQWDRTLIPSH-----TGTSGTPTNIYHGTDPDQVFTIENSVPVHLRTGDEBATGT 403
QY 413 YTFBDVPHSSYAHSSQSLRLMN--PLIDQYLYLSRTNTPS-GITTTQRLQFSQAGASD 469
Db 404 FFFDCKP--CRLTHTWQTNALGLPPLNSLPOSEGATNFGDIGVQDQKRRGVITQMGNTD 461
QY 470 IRDOSRWLPCPCYRQORVSKTSADNNNSEYSGTGAHYHLNCRDLSVNPGB-----A 522
Db 462 YITEATTMRP-----AEVGSAPYSPASTQ-----GPFKTPIAAG 498
QY 523 MASHKDDKEFFPQSGVLIFGKGSSEKTNVDIE-----555
Db 499 RGAQTDENQAADGDPYAFGRHQKQTTTGETPERFYIAHODTGRYPAGDWIQNINF 558
QY 556 KMITDEEERTNPVATQYSGVSTNLQRGNRQAATADVNTQGVLPQMWQDRDVLQ 615
Db 559 NLPTVNDNLLPTDPIG---GKTGINY-----TNFTYGLPTALN-NVPPVYPNG 605
QY 616 PIWAKIPTDGHFHPSPLMGGFGLKHP-----PPQLIKNTVPANPSTTSAKFAS 668
Db 606 QIWDKEFTD-----LKPLHLVNAFPVQNCNCPGQLFVKVAPNLNNEYDPPDASAMSR 658
QY 669 FITQYSTQGVSEIWELEQKENSKRNPETIQYTSNYSNNVD 711
Db 659 IVT-YSDFWKKGLVFKAKLRASHTNP-IQ-----QMSINVD 694

RESULT 10
VCPVIM
coat protein VP1 - minute virus of mice (strain MMWi)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: B23008; B29510
R:Sabli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: B23008
A:Molecule type: DNA
A:Residues: 1-718 <SAH>
A:Cross-references: UNIPROT:P07302; EMBL:X02481; NID:G60918; PIDN:CAB46507.1; PID:G54199
R:Aerall, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, MMV(1), and
A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: B29510
A:Molecule type: DNA
A:Residues: 1-143, 'A', 145-718 <SAH>
A:Cross-references: EMBL:ML2032
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 11.1%; Score 442; DB 1; Length 718;
Best Local Similarity 24.8%; Pred. No. 4.7e-21;

```

```

Matches 201; Conservative 99; Mismatches 294; Indels 218; Gaps 38;
QY 45 LVLPGYKYLGPNGLDGEPVNEADAAALBHDKAYDRQLDSDGNPYLKYNNHADAEPQERL 104
Db 1 MYPFGYKYLGPNGLDGEPNPNFSDAAAKEHDEAYDQYIKSGKNPYLYFSAADQFIDOT 60
QY 105 KEDTSPGNTGRVAFQAKRVLPLGLVBEPPVKTAPGKKRPVEHSPVSDSSGTGKAG- 163
Db 61 KDARDMGKGVGHVFFRTKRAFAPKLATDSEP-----GTSVSRAGK 101
QY 164 -----QOPARKEL-----NFGQT-GRADSVDPDQPPQLGPPAAPSGGLTNTMAT 205
Db 102 RTRPPAYIFINQARAKKLTSAAQQSSQTSMDGTSQPD-----GGGVHSAAR 150
QY 206 -----GSGAPMADNNEGADVGNSSGNWHCDSS--TWMGD-----RVITTSRTWAL-- 249
Db 151 VERAADGPGGSGG-----GGSGGGVGVSVDNTHYRFLGDDGWVEITALATLVLNM 206
QY 250 PTYNN-----HLYKQISSQSGASNDNHFGYSTPGYDFDNRHCHFSRDMQRLINN 303
Db 207 PKSENYCRTRVNTTDSVKGNMAKDDAHEQIWTWPSLVDAWANGVWLQPSDMQYICNTM 266
QY 304 WGRPKRLNFKLNIQVKEVTQND-GTTTIA-----NNLTSTVQVFTDSEYQLPVLSAHO 359
Db 267 SOLNLVSLDQEIFNVVLKTVTEQDSGGOAIKYNDLITACMMVAVDSSNNILPYTPAANS 326
QY 360 GCLPPPPADVFMVPOGY-----LTLNN-----GSAVGRSSPYCYLEYFPS 400
Db 327 ETLGYPWKPTIASPYRYFCVDRDLSVTYENOBGTIEHNVMGTPKGMNSQPTTIENTQQ 386
QY 401 -OMLTGNPF-TFSYTFEDVPFHSSYAHSSQSLRLMN--PLIDQYLYLSRTNTPSGT 456
Db 387 ITLLRTGDEBATGTYYFTNPV--KLTHWTQNRQLGQPELLSTF-----PEADTDAGTLT 440
QY 457 QSRLOFSQAGASDIRDOSRWL-----PGCYRQORVS-----KTSAD-- 494
Db 441 ---AQGRHGATQ---EVNVSEALRTPAQVGFQCPHNDFEASRAGFAAPKVPADVT 494
QY 495 -----NNNSEYSWTGATKYHLNCRDLSVNPQPMASHKDDEKF-----PPOSG 538
Db 495 QGVDRANGSVYSY---GKQH---GENWAAGPAPERYTWDNETNFGSGRDRDGFIOQA 548
QY 539 VLIFGKQSEKTNVLEKMITDEEIRITNPVATEQYSGVSTNLQRGNRQAATADVNTQ 598
Db 549 PLVPPPLNG-----ILTANPDKNDI---HFSNV-----FNSY 581
QY 599 GVLPGMWQDRDYLQGPWAKIPHTDGHFHPSPLMGGFGLKHP-----PQLIKNT 652
Db 582 GLPTAFS-HPSPVYPOGQIWDK-----ELDLHKRPLHITAPFVCKKNAP 625
QY 653 ----VPANPSTTF----SAAKFPASITQYSTQGVSEIWELEQKENSKRNP--EIQVTS 702
Db 626 GQMLVRLGNLTDQVDPNGATLSRIVT-YGTFPFWKGLTMRKLRANNTWNPVYQVSVED 684
QY 703 NYNKSNNVDFTVDTNGVSEPRPIGTRYLTRN 734
Db 685 NGNSYMSVTWKLPTATGNQSVLITRPAEN 716

RESULT 11
VCPVCP
coat protein VP1 - canine parvovirus (strain N)
N:Contains: coat protein VP2
C:Species: canine parvovirus, CPV
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: B29962
R:Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988
A:Title: Nucleotide sequence and genome organization of canine parvovirus.
A:Reference number: A29962; MUID:88062992; PMID:2824850
A:Accession: B29962
A:Molecule type: DNA
A:Residues: 1-748 <REE>

```

A;Cross-references: UNIPROT:P12930; EMBL:M19296

C;Accession: B33302

A;Introns: 26/3

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

F;165-748/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.1%; Score 442; DB 1; Length 748;

Best Local Similarity 24.2%; Pred. No. 5e-21;

Matches 194; Conservative 99; Mismatches 306; Indels 202; Gaps 35;

```

Qy 30 PPPKPAERHK-----DDSR-----GLVLPGYKYLGFPGNLDKGEFVNEADAAALE 74
Db 3 PPPKPAERHKGLVVKWEGKDLITMCFFIGLVPFGYKYLGFPGNLDKGEFVNEADAAALE 62
Qy 75 HDKAYDQDSDGNPILYNHADAEFBERLKEDTSFGNLRGAVFOAKRVLEPL-GLVE 133
Db 63 HDEAYAYLRSGKNPYLYFSPADRFIDQTKADKGGKIGHYFFRAKKAIAVLDTDPD 122
Qy 134 EPVKTAPGKKRPVBHSPVDPSSSGTGK-----AGQOPARKRLNFGQTGDADSDVPDQPL 188
Db 123 HPSTSRPTK--PTKRSPPPHIFINLAKKKAGAGQV---KRDNLAPMSDGAQVQPDG--- 174
Qy 189 GQPPAASGLGTMTATGSG-APMADNNEGADGVNSSGNWH-----CDSTWMDRVI 240
Db 175 GQPAV-----RNERATGSGNGSGGGGGGGVGISTGTNNQTEFKFLENGWV--EIT 226
Qy 241 TTSRTWTALPTYNHLYKQI-----SSQSGASNDNHYFGYSTPMGYDFNFRFCHFS 292
Db 227 ANSRLVHLNMPSENRVRVVVNNMDKTAVNGNMLDDIHAQIVTPSLVLDANAGWVFN 286
Qy 293 PRDQRLINNMWFRPKRLNFKLNIQVKEVTQN---DGTITANNLTSTVQVFTSEYQ 349
Db 287 PGDWQLIVNTMSELHLSVFEQIEINVLKTVSEATQPTKYVNDLTASLWALDSNNT 346
Qy 350 LPYVLGSAHQCL-----PPFPA-----DVFMVPOGYLTINNGSQAVGRS----- 390
Db 347 MPFTPAAMRSETLFGYPMKPTIPTWRYFYFQWDRTLIPSH-----TGTSCTPTNIYHGT 400
Qy 391 ----SFYCLE-YPPSOMLRTGNNTFSYTFEDV-----LMNPLIDQYL--YYLSRTNTPSGTTTQ 457
Db 401 DPDDVQFYTIENSVPVHLRTGDEFAFGTFFDCKPCRLTHWTQNRALGPLFNLSPQ 460
Qy 427 SQSLDR-----LMNPLIDQYL--YYLSRTNTPSGTTTQ 457
Db 461 SEGATNFGDIGVQDQKRGVYTMGNTNVTIETIMRPAEYGSAPYYSFEASTQGPFTK 520
Qy 458 SRLQFSQAGASDIRDQSRNWLPGCYRQORVSKTSADNNSEYSWTGATKYLHNGRSLV 517
Db 521 --IAAGRGGAQTYENQAADGDPYAFGRQHQKTTT-----TGTEPERF----- 562
Qy 518 NPGPAMASHKDDEKEFPQSGVLIFGKGSEKTNVDIEKVMITDDEBEIRTNVATQYQ 577
Db 563 ----TYIAHQDTGR--Yp-----EGOWIQINIF-NLPVTNDNVLLPTDPIG-----G 602
Qy 578 SVSTNLQGRNQRAATADVNTQGLPGLVMVQDRDYLQGPVWAKIPHTDGHFHPSPLMGGF 637
Db 603 KTGINY-----TNIFNTYGLTALN-NVPPVYPNGQIWDKEFTD-----LKPL 646
Qy 638 GLKHP-----PPQIILKNTVPANSTPTSAKFAFSTIQYTGQVSVEIEWELOKEN 690
Db 647 HVNAPFVQNNCPQGLFVKAPNLNTEYDYPDASANMSRIVT-YSDFWWKGKLVFAKLRA 705
Qy 691 SKRWNPETQYTSNYSNVND 711
Db 706 SHTWNP-IQ-----QMSINVD 720

```

RESULT 12

VCPVPP

coat protein VP1 - porcine parvovirus

N;Contains: coat protein VP2

C;Species: porcine parvovirus

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C;Accession: B33302

R;Ranz, A.I.; Mancus, J.J.; Diaz-Aroca, E.; Casal, J.I.

J. Gen. Virol. 70, 2541-2553, 1989

A;Title: Porcine parvovirus: DNA sequence and genome organization.

A;Reference number: A33302; MUID:90010964; PMID:2794971

A;Accession: B33302

A;Molecule type: DNA

A;Residues: 1-723 <RAN>

A;Cross-references: UNIPROT:P18546; EMBL:D00623

C;Genetics:

A;Introns: 10/1

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

F;145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.9%; Score 434.5; DB 1; Length 723;

Best Local Similarity 25.1%; Pred. No. 1.5e-20;

Matches 199; Conservative 107; Mismatches 318; Indels 169; Gaps 37;

```

Qy 30 PPPKPAERHKDDSRGLVLPGYKYLGFPGNLDKGEFVNEADAAALEHDKAYDQDSDGNP 89
Db 3 PPPKPAERHK-----RGTI-----LPGNSLDQGEPTNPSDAAAKEHDEAYDKYKSGNP 49
Qy 90 YLKNHADAEFBERLKEDTSFGNLRGAVFOAKRVLEPLGLVEEPVKTAPGKKRPVHHS 149
Db 50 TFVPSAADERFKETEHAKDYGGKIGHYFFRAKFAFAPKLSETDSPTTS---QQPEVRS 106
Qy 150 PVEPDSSSGTGK-----AGQOPARKRLNFGQTGDADSDVPDQPLGQPPAASGLG 199
Db 107 PRKHGSKPPKRPAPRHIFINLAKKKAGKTSNTNSMSNSEVQHNPI-----NAA 158
Qy 200 TMTATGSG---GAPMADNNEGADGVNSSGNWH---CDSTWMD---RVITSTRTALPT 251
Db 159 TEISATGNSGCGGGGGGAGGVGTVSTGFNQTEFVYLGEGVLVITAHASLIHLNM 218
Qy 252 YNNHLYKQI---SQSGAS---NDNHYFGYSTPMGYDFNFRFCHFSPPRDWQRLNNW 304
Db 219 PEHETVKRIIHLVNSGSGAGQWQDDAHTQWTFWLSLDANAGWVFNPAQWQLISNMT 278
Qy 305 GFRPKRLNFKLNIQVKEVTQN---DGTITANN--LTSTVQVFTSEVQLPYVLSAHQGC 361
Db 279 EINLVSFEQAFINVLKTIITESATSPPTKYVNDLTASLWALDNTNLTPTPAAPRSET 338
Qy 362 LPPPPADVFMVPOGYL-----TLNNGSQAVGRS-----FYCLE-YPPSQ 401
Db 339 LGFYPLPTPTQYRYLVSCIRNLNPPYTGQSQPNRLNRLHSDIMFTIENAVPIH 398
Qy 402 MLRTGNF--TFSYTFEDVPFHSSYAHQSQSLDR-----LMNPLI--DQYLYLSRTNT 450
Db 399 LLRTGDEFSTGIYHFTKPL--KLTHSQWTRSLGLPPKLLTEPTTEGQHPGTLPAANT 456
Qy 451 PSGTTTSRLQFSQAGASDIRDQSRNWLPGCYRQORVSKTSADNNSEYSWTGATKYLH 510
Db 457 RKGHYHOTINNSYTEATA--IRPAQVGY-----PNGAIRFTMDYQHGHLTSSQELERYT 491
Qy 511 NGRDNLNPGPAMASHKDDEKEFPQSGVLIFGKGSEKTNVDIEKVMIT---DDEBEIRT 567
Db 492 NGGPFLLTPVPTADTYNDDE-----PNGAIRFTMDYQHGHLTSSQELERYT 539
Qy 568 TNP-----VATEYQY--SVSTNLQGRNQRAATD-----VNTQGLPGLVMV 606
Db 540 FNPQSKGRAPKQOPNQOAPLNLENTNNGTLLPSDPIGGKSNHMFMTNLTNYGLTALN- 598
Qy 607 QDRDYLQGPVWAKIPHTD--GHFHPSPLMGGFGLK--HPPQIILKNTVPANSTPTSA 663
Db 599 NTAPVFPNGQIWDKELDTLKLRLH---VTAPFYCKNPNPQGLFVKIAP---NLTDFFNA 652
Qy 664 -AKFASPIQYTSQVSVEIEWELOKENSKRWNPETQYTSNYSNVNDVTDVFN--GV- 719
Db 653 DSPQOPRIITYSNFWKGTLTFTAKMRSSNNWNPDIQHTTT---AENIGKIYPTNIGIK 709
Qy 720 ----YSE--RPPI 726

```

Db 710 MFPEYSQLIPRKL 722

RESULT 13

VCPVCD

coat protein VP1 - canine parvovirus (strain CPV-d)

N:Contains: coat protein VP2

C:Species: canine parvovirus, CPV

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: A31163

R:Parish, C.R.; Aquadro, C.F.; Carmichael, L.E.

Viology 166, 293-307, 1988

A:Title: Canine host range and a specific epitope map along with variant sequences in the

A:Reference number: A31163, MUID:89020796; PMID:3176341

A:Accession: A31163

A:Molecule type: DNA

A:Residues: 1-737 <PAR>

A:Cross-references: UNIPROT:P17455; EMBL:M23255; NID:g333467; PIDN:AAA47158.1; PID:g3334

C:introns: 26/3

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

F:584-737/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.7%; Score 425.5; DB 1; Length 737;

Best Local Similarity 24.4%; Pred. No. 5.9e-20;

Matches 190; Conservative 102; Mismatches 317; Indels 171; Gaps 34;

QY 30 PPPPAERHKDD----SRGLVLPGVKYLGPNGLIDKGEPVNEADAAALEHDKAYDRQLDS 85

DB 3 PPAAARRGKGVLVKWEKGKDLITYKYLGPNSLIDQCEPTNPSPDAAAKEHDEAAVAYLRS 62

QY 86 GDNPLYKVNHADAEFBLKEDTSPGGNLGRAVFAQAKRVLEPL-GLVEEPPVKTAPGKKR 144

DB 63 GKNPYLFSPADQRFIDQTKDAKDWGKIGHYFFRAKKAIAPVLTDPDHPSTSRPTK-- 120

QY 145 PVEHSPVEPPSSSGTGK----AGQAPARKLNFGQTGDADSDVPDPOLGQPPAAPSLGL 199

DB 121 PTKRSKPPPHIFINLAKKKKAGAGV---KRDNLAPMSDGAQVDPG---GQPAV----- 168

QY 200 TYNMATGSG-APMADNNEGADGVNSSGNWH-----CDSTWMDRVITTSRTWALPT 251

DB 169 RNERATYSGNGSGGGGGGGGGVIGTGTENNQTPEKFLENGWV--EITANSRLVHLNM 226

QY 252 YNNHLYKQI-----SSQSGASNDNHVEGYSTPWGYFDNFRPHCHFSPRDQWRLLNN 303

DB 227 PSENYRRVVVNMDKTAVNGNMLADODIHAQIVTPMWSLVANAGWYFNPGDWQLIVNTM 286

QY 304 WGFPRKRLNFKLPIQVKEVTQN---DGTGTTANNLTSTVQVETDSEYQLPYVLGSAHQ 360

DB 287 SELHLVSPQEIEFNVVLKTVSBSATQPTKVYNNDLTASLWVALDSNNTPFTPAAMRSE 346

QY 361 CL-----PPFFA-----DVFWVPQGYLTLLNGSQAVRS-----SPCYCLE 396

DB 347 TLGFYPMKEPTITPWRYYFQWDRTLPSH-----TGTSGTPTNIYHGTDPDDVQFYTIE 400

QY 397 -YFPQMRLTGNPF-TFSYTFEDVPFHHSSYAHSSQSLDRMLNPLIDQLYLXSRNTPTSGT 454

DB 401 NSVPVHLLRTGDBFATGTFFFDCKP--CRLTHTWQTNRAIG--LPPFLNSLPQS---EGA 453

QY 455 TTQSRLOPFSQAGASDIRQSRNWLPGCYRQORVSKTSADNNNSE--YSWTGATKYHLNG 512

DB 454 T-----NFGDIGVQDQKRGVYTMQNTYITTEATIMRPAEYGVSAFYYSFEASTQ----- 503

QY 513 RDSLVNPGP-----AMASHKDDEKFFPSGVLIFGQKSGSEKTNVDIE----- 555

DB 504 -----GPFKPTIAAGRGGAQTDENQAADGNPRYAFGRQHGQKTTTITGETPERFYIAH 556

QY 556 -----KVMITDSEETIRTNPVATEQYGSVSTNLQRNQRAATADVNTQ 598

DB 557 QDTCRYPEGDWIQINFLPVTNDVLLPTDPIG-----GKTGINY-----TNINFTY 604

```
QY 556 -----KVMITDEEIRITNPVATEQYGSYSTNLQRNQAAATADVNTQGVLPQMV 605
D 549 EGDWQINFNLPVTDNVLPIPIG-----GKTGINY-----TNIFNTYGPLTALN 596
QY 606 WQDRDVLQYQIPWAKIPIHTDGHFHPSPLMGGFGLKHP-----PPQILIKNTVPANPS 658
D 597 -NVPVPVTPNGQIWDKBEPTD-----LKPLHLVNAFPVCONNCPGQLFVKLAENLTNEY 648
QY 659 TTFSAAKFASPIQYSTQGVSEIWELOKENSKRNPFIQYTSYNKSVNVD 711
D 649 DPDASANNRIVT-YSDPFWKGLVFKAKLRASHTNP-IQ-----QMSINVD 694

RESULT 15
VCPVV2
coat protein VP1 - parvovirus H1
C:Species: parvovirus H1
A:Note: host Homo sapiens (man)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
R:Accession: A03699
R:Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A:Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization
A:Reference number: A03695; NUID:83112183; PMID:6823009
A:Accession: A03699
A:Molecule type: DNA
A:Residues: 1-722 <RHO>
A:Cross-references: UNIPROT:P03136; EMBL:X01457; EMBL:J02198
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match          9.9%; Score 397; DB 1; Length 722;
Best Local Similarity 23.7%; Pred. No. 4.2e-18;
Matches 187; Conservative 110; Mismatches 325; Indels 168; Gaps 37;

QY 30 PPPKPAERHDDSRGLVLPQYKYLGPENGLDKGPFVNEADAAALEHDKAYDRQLDSDGNP 89
D 3 PPAKRAKRG-----NSLDQGEPTNPSSDAAAKEHDEAYDQYIKSGKNP 44

QY 90 YLKNYHADAEFOERLKEDTSFGNIGRAVQAKKRVLEPLGLVEEPVKTA---EGKK-RP 145
D 45 YLYFSPADQRFIDQTKAKOMGGKGHYFFRTKRAFAPKLSTDSEPTGSGVSRFGKRTKP 104

QY 146 VEHSVPEDSSSGTGKAGQAPARKLNFGTGDADSVDPDPLGPPAAPSGGLTNTMAT 205
D 105 PAHIFVQ-----ARAKKASLAQAQRTLTMSDGTETNPQDTGIANARVERSAD 154

QY 206 GSGAPMADNNEGADVGNSSGNHCDSTW--MGDRVITTTSTRT-----WALPTYNNHLYK 258
D 155 GGGG-SGGGGGGGGIGVSTGYDNTQTYKFLGDGWVEITAHASRLHLGLMPPSENICRV 213

QY 259 QISSQSGASNDNHVFG---YS-----TPWGYDFPNRPHCFSPRDWQRLNNNWGFRPK 309
D 214 TVHNNQTTGHGTVKYGNAVDTHQOIWTPNSLVDANAGVWFQPSDWQFIQNSMESLNLD 273

QY 310 RLNFELNIOVKEVTQNDGTTTIA-----NNLTSTVOVFTDSEYQLPYVLGSAHQGCL-- 362
D 274 SLSELENNVVKVTTEQOGAGQDAIKVYNNDLTACMMVALDSNNILPYTPAAQTSSETLGP 333

QY 363 -----PPFPAD---VFMVPOYGYLTNNQSQ-----AVG-----RSSFYCLE-YPPSQML 403
D 334 YPWKPTAPAPVRYFFMPRQLSVTSSNSAEGTQITDTIGEPQALNSQFFTIENLTLPITLL 393

QY 404 RTGNNTTF-SYTFEDVPHSSYA-----HSQSLRLMN-PLIDQYLYLSRTNTPSGTT- 455
D 394 RTGDEFTTGYIENFTPLKLTHTWQTRHLACLQGITDLEPTSDTATASLTANGDRFGSTQ 453

QY 456 TQS-----RLQFSQAGASDIRDQSRNMLPGPCYQQRVSKTSAD-NNNSEYSWTGAT 506
D 454 TQNNVYTEALRTPAQIGFMQPHDNFEANGGPF-----KPVVPLDITAGEDHDANGAI 509

QY 507 KYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVLI FGQSGEKTNDVIEKVMITDEBEIR 566
```

```
Db 510 RFNYG-----KQHGED-----WAKQGAAPERYTWDAI---: : : : :
QY 567 TT-----NPVATEQYGSYSTN-----LQRGNROAATADV-----NTQGVLPQMVWQDR 609
D 543 DTARCFVQSAPI-----SIPPNQOILQREDAIAGRTNMHYTNVFNISYGPLSAPPHPD- 595
QY 610 DVYLOQPIWAK---IPHTDGHFHPSPLMGGFGLKHPPIQLI---KNTVPANPSTTFSA 663
D 596 PIYPNGQIWDKELDLHKPRLHVTAPFV---CKNPPGQLFVHLGNLTDQFDENST--- 649
QY 664 AKFASPIQYSTQGVSEIWELOKENSKRNPFIQYTSN--YKSVNVD--FTVDITNGV 719
D 650 --TVSRIVTYSTFYWKGIKFKAKLRPNLTWNPVYQATTDTSVANSYMNKKWLPSATGNM 707

QY 720 YSEP---RPI 726
D 708 HSDPLICRPV 717
```

Search completed: May 19, 2005, 13:33:49
Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:33:06 ; Search time 139 Seconds
(without alignments)
1768.799 Million cell updates/sec

Title: US-10-038-972A-13

Perfect score: 3994

Sequence: 1 MAADGYLPDWLEDTLSGIR.....TNGVSEPRPIGTRYLTRLN 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3994	100.0	735	13	US-10-038-972A-13
2	3994	100.0	735	14	US-10-293-478-1
3	3994	100.0	735	14	US-10-291-583-70
4	3994	100.0	735	15	US-10-423-704A-4
5	3994	100.0	735	17	US-10-496-799-4
6	3989	99.9	735	17	US-10-880-297-11
7	3985	99.8	735	17	US-10-880-297-15
8	3984	99.7	735	17	US-10-880-297-5
9	3984	99.7	735	17	US-10-880-297-13
10	3984	99.7	735	17	US-10-880-297-19
11	3983	99.7	735	17	US-10-880-297-7
12	3980	99.6	735	17	US-10-880-297-9
13	3980	99.6	735	17	US-10-880-297-31

14	3975	99.5	735	17	US-10-880-297-25
15	3975	99.5	735	17	US-10-880-297-27
16	3969	99.4	735	17	US-10-880-297-17
17	3969	99.4	735	17	US-10-880-297-39
18	3962	99.2	735	17	US-10-880-297-23
19	3962	99.2	735	17	US-10-880-297-41
20	3944	98.7	735	17	US-10-880-297-35
21	3941	98.7	735	17	US-10-880-297-21
22	3936	98.5	735	17	US-10-880-297-37
23	3930	98.4	735	17	US-10-880-297-29
24	3929	98.4	735	17	US-10-880-297-33
25	3623	90.7	735	14	US-10-291-583-67
26	3622	90.7	735	14	US-10-291-583-69
27	3617	90.6	735	14	US-10-291-583-68
28	3615	90.5	735	14	US-10-291-583-66
29	3532.5	88.4	736	14	US-10-291-583-71
30	3532.5	88.4	736	15	US-10-423-704A-6
31	3532.5	88.4	736	17	US-10-496-799-6
32	3421	85.7	731	14	US-10-291-583-88
33	3416.5	85.5	738	14	US-10-291-583-91
34	3409.5	85.4	736	14	US-10-291-583-65
35	3405.5	85.3	738	14	US-10-291-583-93
36	3405.5	85.3	738	14	US-10-291-583-94
37	3404.5	85.2	738	14	US-10-291-583-79
38	3404.5	85.2	738	14	US-10-291-583-81
39	3404.5	85.2	736	14	US-10-291-583-85
40	3402.5	85.2	736	14	US-10-291-583-64
41	3402.5	85.2	736	15	US-10-423-704A-5
42	3402.5	85.2	736	15	US-10-696-261-3
43	3402.5	85.2	736	15	US-10-696-261-13
44	3402.5	85.2	736	15	US-10-696-282-3
45	3402.5	85.2	736	15	US-10-696-282-13

ALIGNMENTS

[Handwritten signature]

RESULT 1
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Publication No. US20020192823A1
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VLP capsid protien
; US-10-038-972A-13

Query Match	100.0%;	Score	3994;	DB	13;	Length	735;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	735;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
Qy	1	MAADGYLPDWLEDTLSGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60				
Db	1	MAADGYLPDWLEDTLSGIRQWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60				
Qy	61	KGEVPEVNEADAALAEHDKAYDRQLDSGDNPNYLYKNHADAFOERLKEDTSGFNGNLGRAVFQ	120				
Db	61	KGEVPEVNEADAALAEHDKAYDRQLDSGDNPNYLYKNHADAFOERLKEDTSGFNGNLGRAVFQ	120				
Qy	121	AKKRVLEPLGLIVEPVTAPGKKPVEHSPVEPSSSTGKAGQOPARKRLNFGQTGDAD	180				
Db	121	AKKRVLEPLGLIVEPVTAPGKKPVEHSPVEPSSSTGKAGQOPARKRLNFGQTGDAD	180				

181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVNCGNSGNWHCDSTWMDRVI 240
Db
181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVNCGNSGNWHCDSTWMDRVI 240
Qy
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWPYDFNRFCHFSPRDWORLI 300
Db
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWPYDFNRFCHFSPRDWORLI 300
Qy
301 NNNWGFPRKLNFKLNFQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
Db
301 NNNWGFPRKLNFKLNFQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
Qy
361 CLPPFPADVFWVPOYGYLTLLNGSQAVERSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Db
361 CLPPFPADVFWVPOYGYLTLLNGSQAVERSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Qy
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Db
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Qy
481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
Db
481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
Qy
541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Db
541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Qy
601 LPGMWQDRDYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db
601 LPGMWQDRDYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy
661 FSAAKFASFIQYSTQGVSVIEWELQENSKRNPEIQYTSNKNKSNVNDFTVDNTGVY 720
Db
661 FSAAKFASFIQYSTQGVSVIEWELQENSKRNPEIQYTSNKNKSNVNDFTVDNTGVY 720
Qy
721 SEPRPIGTRYLTRNL 735
Db
721 SEPRPIGTRYLTRNL 735

RESULT 2

US-10-293-478-1
; Sequence 1, Application US/10293478
; Publication No. US20030078411A1
; GENERAL INFORMATION:
; APPLICANT: PATEL, SATIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; PRIOR FILING DATE: 2002-11-14
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 100.0%; Score 3994; DB 14; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAADGYPDLWLEDTLSGIRQWKLKPGPPPPKPAERHKDDSRGLVLPKYKYLGPENGLD 60
Db 1 MAADGYPDLWLEDTLSGIRQWKLKPGPPPPKPAERHKDDSRGLVLPKYKYLGPENGLD 60
Qy 61 KGPVNEADAAALEHDKAYDRQLDSGDNVLYKYNHADAFOERLKBEDTSFGNGLGRAVPQ 120
Db 61 KGPVNEADAAALEHDKAYDRQLDSGDNVLYKYNHADAFOERLKBEDTSFGNGLGRAVPQ 120

121 AKKRVLEPLGLVEBPVKTAQKPKRPEVHSVPEVPSDSSGTGKAGQOPARKLNFGQTGDAD 180
Db
121 AKKRVLEPLGLVEBPVKTAQKPKRPEVHSVPEVPSDSSGTGKAGQOPARKLNFGQTGDAD 180
Qy
181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVNCGNSGNWHCDSTWMDRVI 240
Db
181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVNCGNSGNWHCDSTWMDRVI 240
Qy
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWPYDFNRFCHFSPRDWORLI 300
Db
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWPYDFNRFCHFSPRDWORLI 300
Qy
301 NNNWGFPRKLNFKLNFQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
Db
301 NNNWGFPRKLNFKLNFQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
Qy
361 CLPPFPADVFWVPOYGYLTLLNGSQAVERSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Db
361 CLPPFPADVFWVPOYGYLTLLNGSQAVERSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Qy
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Db
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Qy
481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
Db
481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
Qy
541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Db
541 IFGQKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Qy
601 LPGMWQDRDYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db
601 LPGMWQDRDYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy
661 FSAAKFASFIQYSTQGVSVIEWELQENSKRNPEIQYTSNKNKSNVNDFTVDNTGVY 720
Db
661 FSAAKFASFIQYSTQGVSVIEWELQENSKRNPEIQYTSNKNKSNVNDFTVDNTGVY 720
Qy
721 SEPRPIGTRYLTRNL 735
Db
721 SEPRPIGTRYLTRNL 735

RESULT 3

US-10-291-583-70
; Sequence 70, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (AAV) Sequences
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US-60/350,607
; PRIOR FILING DATE: 2001-11-12
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV2

US-10-291-583-70

```
Query Match      100.0%; Score 3994; DB 14; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPPNGLD 60
Db 1 MAADGVLDPWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPPNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120
Db 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120

QY 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKGAGQQPARKRLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKGAGQQPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240

QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYDFDNRFCHFSPRDWORLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYDFDNRFCHFSPRDWORLI 300

QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360

QY 361 CLPPPADVFMVPOGYLTLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVPF 420
Db 361 CLPPPADVFMVPOGYLTLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVPF 420

QY 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQPSQAGASDIRQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQPSQAGASDIRQSRNWLPG 480

QY 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540

QY 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600

QY 601 LPGMWQDRDYYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPQILIKNTVPANPSTT 660
Db 601 LPGMWQDRDYYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPQILIKNTVPANPSTT 660

QY 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

RESULT 4

```
US-10-423-704A-4
; Sequence 4, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
```

RESULT 5

```

; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PR1
; ORGANISM: adeno-associated virus serotype 2
US-10-423-704A-4

Query Match      100.0%; Score 3994; DB 15; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPPNGLD 60
Db 1 MAADGVLDPWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPPNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120
Db 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120

QY 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKGAGQQPARKRLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKGAGQQPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240

QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYDFDNRFCHFSPRDWORLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYDFDNRFCHFSPRDWORLI 300

QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360

QY 361 CLPPPADVFMVPOGYLTLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVPF 420
Db 361 CLPPPADVFMVPOGYLTLNNGSAVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVPF 420

QY 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQPSQAGASDIRQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQPSQAGASDIRQSRNWLPG 480

QY 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540

QY 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQGRNQAAATADVNTQGV 600

QY 601 LPGMWQDRDYYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPQILIKNTVPANPSTT 660
Db 601 LPGMWQDRDYYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPQILIKNTVPANPSTT 660

QY 661 FSAAKFASFIQYSTQGVSVIEWELQKNSKRWNPFIQYTSNKNKSVNVDFTVDINGVY 720
Db 661 FSAAKFASFIQYSTQGVSVIEWELQKNSKRWNPFIQYTSNKNKSVNVDFTVDINGVY 720

QY 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

US-10-496-799-4
 ; Sequence 4, Application US/10496799
 ; Publication No. US20050014262A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Trustees of The University of Pennsylvania
 ; APPLICANT: Geo, Guangping
 ; APPLICANT: Wilson, James M.
 ; APPLICANT: Alvira, Mauricio
 ; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 9 Sequences, Vectors Contain
 ; FILE OF INVENTION: Same, and Uses Therefor
 ; FILE REFERENCE: UPN-02734FCI
 ; CURRENT APPLICATION NUMBER: US/10/496,799
 ; PRIOR FILING DATE: 2004-06-08
 ; PRIOR FILING DATE: 2001-12-17
 ; PRIOR APPLICATION NUMBER: US 60/341,150
 ; PRIOR FILING DATE: 2002-06-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 735
 ; TYPE: PRT
 ; ORGANISM: adeno-associated virus serotype 2
 US-10-496-799-4

Query Match
 Best Local Similarity 100.0%; Score 3994; DB 17; Length 735;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD	60
DB	1	MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD	60
QY	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPKYKYNHADADEFQRLKEDTSFGNIGRAVQ	120
DB	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPKYKYNHADADEFQRLKEDTSFGNIGRAVQ	120
QY	121	AKKRVLEPLGLVEEPVKTAQKGRPVESHSPVPSDSSGTGKAGQAPARKLNFGQTGDAD	180
DB	121	AKKRVLEPLGLVEEPVKTAQKGRPVESHSPVPSDSSGTGKAGQAPARKLNFGQTGDAD	180
QY	181	SVPDPQLGPPAAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI	240
DB	181	SVPDPQLGPPAAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI	240
QY	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDNRFCHFSPRDWORLI	300
DB	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDNRFCHFSPRDWORLI	300
QY	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ	360
DB	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ	360
QY	361	CLPPPPADVFMVPOGYLTLLNGSQAGRSFICYLFPPSQMLRTGNNTFTSYTFEDVFP	420
DB	361	CLPPPPADVFMVPOGYLTLLNGSQAGRSFICYLFPPSQMLRTGNNTFTSYTFEDVFP	420
QY	421	HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG	480
DB	421	HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG	480
QY	481	PCYRQQRVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEKFFPQSGVL	540
DB	481	PCYRQQRVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEKFFPQSGVL	540
QY	541	IFGQSGSEKTNVDIEKVMITDEBEIRTTNPVATPEQYGSVSTNLQRGNQAATADVNTQGV	600
DB	541	IFGQSGSEKTNVDIEKVMITDEBEIRTTNPVATPEQYGSVSTNLQRGNQAATADVNTQGV	600
QY	601	LPGMWQDRDVLQGPFWAKIPTHGCHFSPLMGFGFLKHPPOILLKNTVPANPSTT	660
DB	601	LPGMWQDRDVLQGPFWAKIPTHGCHFSPLMGFGFLKHPPOILLKNTVPANPSTT	660

QY	661	FSAAKFASFTQYSTGQSVSEIWELOKENSKRWNPEIQYTSNYSKSNVVDFTVDINGVY	720
DB	661	FSAAKFASFTQYSTGQSVSEIWELOKENSKRWNPEIQYTSNYSKSNVVDFTVDINGVY	720
QY	721	SEPRPIGTRYLTRNL	735
DB	721	SEPRPIGTRYLTRNL	735

RESULT 6
 US-10-880-297-11
 ; Sequence 11, Application US/10880297
 ; Publication No. US20050053922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SCHAFER, DAVID V.
 ; APPLICANT: KASPAR, BRIAN
 ; APPLICANT: MAESTRI, NARENDRA
 ; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
 ; FILE REFERENCE: BERK-024
 ; CURRENT APPLICATION NUMBER: US/10/880,297
 ; PRIOR FILING DATE: 2004-06-28
 ; PRIOR FILING DATE: 2003-06-30
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 735
 ; TYPE: PRT
 ; ORGANISM: adeno-associated virus-2
 US-10-880-297-11

Query Match
 Best Local Similarity 99.9%; Score 3989; DB 17; Length 735;
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD	60
DB	1	MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD	60
QY	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPKYKYNHADADEFQRLKEDTSFGNIGRAVQ	120
DB	61	KGEPVNEADAAALEHDKAYDRQLDSGDNPKYKYNHADADEFQRLKEDTSFGNIGRAVQ	120
QY	121	AKKRVLEPLGLVEEPVKTAQKGRPVESHSPVPSDSSGTGKAGQAPARKLNFGQTGDAD	180
DB	121	AKKRVLEPLGLVEEPVKTAQKGRPVESHSPVPSDSSGTGKAGQAPARKLNFGQTGDAD	180
QY	181	SVPDPQLGPPAAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI	240
DB	181	SVPDPQLGPPAAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI	240
QY	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDNRFCHFSPRDWORLI	300
DB	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDNRFCHFSPRDWORLI	300
QY	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ	360
DB	301	NNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ	360
QY	361	CLPPPPADVFMVPOGYLTLLNGSQAGRSFICYLFPPSQMLRTGNNTFTSYTFEDVFP	420
DB	361	CLPPPPADVFMVPOGYLTLLNGSQAGRSFICYLFPPSQMLRTGNNTFTSYTFEDVFP	420
QY	421	HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG	480
DB	421	HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG	480
QY	481	PCYRQQRVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEKFFPQSGVL	540
DB	481	PCYRQQRVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDEKFFPQSGVL	540
QY	541	IFGQSGSEKTNVDIEKVMITDEBEIRTTNPVATPEQYGSVSTNLQRGNQAATADVNTQGV	600

```
Db 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNQAAATADVNTQGV 600
Qy 601 LPGVMWQDRDVLQGPPIWAKIPHDTDGHPHPSPLMGFGGLKHPPPQILLIKNTVPANPSTT 660
Db 601 LPGVMWQDRDVLQGPPIWAKIPHDTDGHPHPSPLMGFGGLKHPPPQILLIKNTVPANPSTT 660
Qy 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Db 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 7
US-10-880-297-15
; Sequence 15, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-15

Query Match 99.8%; Score 3985; DB 17; Length 735;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Qy 61 KGEVNEADAALHDKAYDRQLDSDGNPILKYNHADADEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAALHDKAYDRQLDSDGNPILKYNHADADEFQERLKEDTSFGNLRGAVFQ 120
Qy 121 AKKRVLEPLGLVEEPVKTAPGKRPRVHSVPEDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPRVHSVPEDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Qy 181 SVDPDQLQGPAPPSGLGTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240
Db 181 SVDPDQLQGPAPPSGLGTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240
Qy 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYPDFNRFCHFSPRDQRLI 300
Db 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYPDFNRFCHFSPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLGSAHQG 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLGSAHQG 360
Qy 421 HSSVAHSQSLDRLMNLPLIDQYLYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNLPLIDQYLYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
```

```
Db 421 HSSVAHSQSLDRLMNLPLIDQYLYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Qy 481 PCYRQORVSKTSADNNNSEYSWTKATKYHLNAGRDSLVPNPGPAMASHKDDDEKFFPQSGVYL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTKATKYHLNAGRDSLVPNPGPAMASHKDDDEKFFPQSGVYL 540
Qy 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNQAAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNQAAATADVNTQGV 600
Qy 601 LPGVMWQDRDVLQGPPIWAKIPHDTDGHPHPSPLMGFGGLKHPPPQILLIKNTVPANPSTT 660
Db 601 LPGVMWQDRDVLQGPPIWAKIPHDTDGHPHPSPLMGFGGLKHPPPQILLIKNTVPANPSTT 660
Qy 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Db 661 FSAAKFASFIQYSGTQGVSVIEWELQKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 8
US-10-880-297-5
; Sequence 5, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 733
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-5

Query Match 99.7%; Score 3984; DB 17; Length 733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Qy 61 KGEVNEADAALHDKAYDRQLDSDGNPILKYNHADADEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAALHDKAYDRQLDSDGNPILKYNHADADEFQERLKEDTSFGNLRGAVFQ 120
Qy 121 AKKRVLEPLGLVEEPVKTAPGKRPRVHSVPEDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPRVHSVPEDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Qy 181 SVDPDQLQGPAPPSGLGTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240
Db 181 SVDPDQLQGPAPPSGLGTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240
Qy 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYPDFNRFCHFSPRDQRLI 300
Db 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYPDFNRFCHFSPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLGSAHQG 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLGSAHQG 360
```

QY 361 CLPPFPADVPMVPOYGYLTLNNGSQAVGRSSFYCLFYPSQMLRTGNNFTSYTFEDVPF 420
 Db 361 CLPPFPADVPMVPOYGYLTLNNGSQAVGRSSFYCLFYPSQMLRTGNNFTSYTFEDVPF 420
 QY 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTITTSQRLQFSQAGASDIRDQSRNWLPG 480
 Db 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTITTSQRLQFSQAGASDIRDQSRNWLPG 480
 QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDDEKFPQSGVL 540
 Db 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDDEKFPQSGVL 540
 QY 541 IFKGQSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
 Db 541 IFKGQSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
 QY 601 LFGMWQDRDVLQGPIMAKIPHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
 Db 601 LFGMWQDRDVLQGPIMAKIPHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
 QY 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
 Db 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
 QY 721 SEPRPIGTRYLTR 733
 Db 721 SEPRPIGTRYLTR 733

RESULT 9

US-10-880-297-13

; Sequence 13, Application US/10880297

; Publication No. US20050053922A1

; GENERAL INFORMATION:

; APPLICANT: SCHAFER, DAVID V.

; APPLICANT: KASPAR, BRIAN

; APPLICANT: MAHESRI, NARENDRA

; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS

; FILE REFERENCE: BERK-024

; CURRENT APPLICATION NUMBER: US/10/880,297

; PRIOR FILING DATE: 2004-06-28

; PRIOR FILING DATE: 2003-06-30

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 735

; TYPE: PRT

; ORGANISM: adeno-associated virus-2

US-10-880-297-13

Query Match

Best Local Similarity 99.7%; Score 3984; DB 17; Length 735;

Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSLGLVLPKYGLGPNGLD 60

Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSLGLVLPKYGLGPNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPNLYKYNHADAERLKEKDTSGGNLGRAVFQ 120

Db 61 KGEVNEADAALAHDKAYDRQLDSDGNPNLYKYNHADAERLKEKDTSGGNLGRAVFQ 120

QY 121 AKKRVLEPLGLVEPVKTAQKRPVHSPVPEPSSSGTGKAGQAPARKRNFQGTGDAD 180

Db 121 AKKRVLEPLGLVEPVKTAQKRPVHSPVPEPSSSGTGKAGQAPARKRNFQGTGDAD 180

QY 181 SVDPDPLGQPPAAPSGLTNTMATSGAPMADNNEGADGVNNGNHCDSWMDRVI 240

Db 181 SVDPDPLGQPPAAPSGLTNTMATSGAPMADNNEGADGVNNGNHCDSWMDRVI 240

QY 241 TTSTRTALPTYNHLYKQISSQSGASNDNHYFCYSTPWGYFDNPRFCHFSPRDWORLI 300
 Db 241 TTSTRTALPTYNHLYKQISSQSGASNDNHYFCYSTPWGYFDNPRFCHFSPRDWORLI 300
 QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTIANNTSTVQVFTDSEYQLPYVLGSAHQ 360
 Db 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTIANNTSTVQVFTDSEYQLPYVLGSAHQ 360
 QY 361 CLPPFPADVPMVPOYGYLTLNNGSQAVGRSSFYCLFYPSQMLRTGNNFTSYTFEDVPF 420
 Db 361 CLPPFPADVPMVPOYGYLTLNNGSQAVGRSSFYCLFYPSQMLRTGNNFTSYTFEDVPF 420
 QY 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTITTSQRLQFSQAGASDIRDQSRNWLPG 480
 Db 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTITTSQRLQFSQAGASDIRDQSRNWLPG 480
 QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDDEKFPQSGVL 540
 Db 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDDEKFPQSGVL 540
 QY 541 IFKGQSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
 Db 541 IFKGQSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
 QY 601 LFGMWQDRDVLQGPIMAKIPHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
 Db 601 LFGMWQDRDVLQGPIMAKIPHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
 QY 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
 Db 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
 QY 721 SEPRPIGTRYLTR 735
 Db 721 SEPRPIGTRYLTR 735

RESULT 10

US-10-880-297-19

; Sequence 19, Application US/10880297

; Publication No. US20050053922A1

; GENERAL INFORMATION:

; APPLICANT: SCHAFER, DAVID V.

; APPLICANT: KASPAR, BRIAN

; APPLICANT: MAHESRI, NARENDRA

; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS

; FILE REFERENCE: BERK-024

; CURRENT APPLICATION NUMBER: US/10/880,297

; PRIOR FILING DATE: 2004-06-28

; PRIOR FILING DATE: 2003-06-30

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 19

; LENGTH: 735

; TYPE: PRT

; ORGANISM: adeno-associated virus-2

US-10-880-297-19

Query Match

Best Local Similarity 99.7%; Score 3984; DB 17; Length 735;

Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSLGLVLPKYGLGPNGLD 60

Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSLGLVLPKYGLGPNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPNLYKYNHADAERLKEKDTSGGNLGRAVFQ 120

Db 61 KGEVNEADAALAHDKAYDRQLDSDGNPNLYKYNHADAERLKEKDTSGGNLGRAVFQ 120

QY 121 AKKRVLEPLGLVEPVKTAQKRPVHSPVPEPSSSGTGKAGQAPARKRNFQGTGDAD 180

```
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPSSSGTGKAGQQPARKRLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Db 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Qy 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGTFDNRHCHFSFPRDQRLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGTFDNRHCHFSFPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVFWVPOYGYLTLLNNGSOAVGRSSFYCLEYFPPSOMLRTGNNFTFSTYTEDVVF 420
Db 361 CLPPFPADVFWVPOYGYLTLLNNGSOAVGRSSFYCLEYFPPSOMLRTGNNFTFSTYTEDVVF 420
Qy 421 HSSVAHSQSLDRLMNLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNLPG 480
Db 421 HSSVAHSQSLDRLMNLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNLPG 480
Qy 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Qy 541 IFGKGSEKTNVDIEKWMITDEEIRTTNPVATEQVGSVSTNLQSGNQAATADVNTQGV 600
Db 541 IFGKGSEKTNVDIEKWMITDEEIRTTNPVATEQVGSVSTNLQSGNQAATADVNTQGV 600
Qy 601 LPMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db 601 LPMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSQRWNPETQYTSNKNYSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSQRWNPETQYTSNKNYSVNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

RESULT 11

```
US-10-880-297-7
; Sequence 7, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-7
```

```
Query Match 99.7%; Score 3983; DB 17; Length 735;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 MAADGYPDLWLEDTLSEGIQWKKLPKPPPPKPAERKDDSRGLVLPKYKYLGPNGLD 60
|||||
```

```
Db 1 MAADGYPDLWLEDTLSEGIQWKKLPKPPPPKPAERKDDSRGLVLPKYKYLGPNGLD 60
Qy 61 KGEVYNEDAAALEHDKAYDROLDSDGNPILKYNHADADEFQERLKEDTSFGNGLGRAVFO 120
Db 61 KGEVYNEDAAALEHDKAYDROLDSDGNPILKYNHADADEFQERLKEDTSFGNGLGRAVFO 120
Qy 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPSSSGTGKAGQQPARKRLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPSSSGTGKAGQQPARKRLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Db 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Qy 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGTFDNRHCHFSFPRDQRLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGTFDNRHCHFSFPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVFWVPOYGYLTLLNNGSOAVGRSSFYCLEYFPPSOMLRTGNNFTFSTYTEDVVF 420
Db 361 CLPPFPADVFWVPOYGYLTLLNNGSOAVGRSSFYCLEYFPPSOMLRTGNNFTFSTYTEDVVF 420
Qy 421 HSSVAHSQSLDRLMNLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNLPG 480
Db 421 HSSVAHSQSLDRLMNLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNLPG 480
Qy 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Qy 541 IFGKGSEKTNVDIEKWMITDEEIRTTNPVATEQVGSVSTNLQSGNQAATADVNTQGV 600
Db 541 IFGKGSEKTNVDIEKWMITDEEIRTTNPVATEQVGSVSTNLQSGNQAATADVNTQGV 600
Qy 601 LPMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db 601 LPMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSQRWNPETQYTSNKNYSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSQRWNPETQYTSNKNYSVNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

RESULT 12

```
US-10-880-297-9
; Sequence 9, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; CURRENT FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-9
```

```
Query Match      99.6%; Score 3980; DB 17; Length 735;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy. 1 MAADGYLPDLEDTLSGIRQWMLKPGPPPPKPAERHKDDSGRLVLPYKYLGPNGLD 60
Db 1 MAADGYLPDLEDTLSGIRQWMLKPGPPPPKPAERHKDDSGRLVLPYKYLGPNGLD 60

Qy 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120

Qy 121 AKKRVLEPLGLVEEPVKTPAGKKRPVEHSPVSPDSSSGTGKAGQQPARKRLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTPAGKKRPVEHSPVSPDSSSGTGKAGQQPARKRLNFGQTGDAD 180

Qy 181 SVDPDPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Db 181 SVDPDPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240

Qy 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSPRDWQRLI 300
Db 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSPRDWQRLI 300

Qy 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360
Db 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360

Qy 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPPSQMLRTGNNFTSYTFEDLPF 420
Db 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPPSQMLRTGNNFTSYTFEDLPF 420

Qy 421 HSSYAHQSGLDLMLNPLIDQYLYLSRNTTSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSYAHQSGLDLMLNPLIDQYLYLSRNTTSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480

Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540

Qy 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600
Db 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600

Qy 601 LFGMWQDRDVLQGPWQIWKIPHTDGHFHPSPLMGFGFLKHPPIQLIKNTVPANPSTT 660
Db 601 LFGMWQDRDVLQGPWQIWKIPHTDGHFHPSPLMGFGFLKHPPIQLIKNTVPANPSTT 660

Qy 661 FSAAKFASFIQYSTGQSVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFTVDNGVY 720
Db 661 FSAAKFASFIQYSTGQSVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFTVDNGVY 720
```

RESULT 13

```
US-10-880-297-31
; Sequence 31, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAESTRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; FILE REFERENCES: BERK-024
; CURRENT FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
```

```
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PaatSEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-31

Query Match      99.6%; Score 3980; DB 17; Length 735;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAADGYLPDLEDTLSGIRQWMLKPGPPPPKPAERHKDDSGRLVLPYKYLGPNGLD 60
Db 1 MAADGYLPDLEDTLSGIRQWMLKPGPPPPKPAERHKDDSGRLVLPYKYLGPNGLD 60

Qy 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120

Qy 121 AKKRVLEPLGLVEEPVKTPAGKKRPVEHSPVSPDSSSGTGKAGQQPARKRLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTPAGKKRPVEHSPVSPDSSSGTGKAGQQPARKRLNFGQTGDAD 180

Qy 181 SVDPDPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Db 181 SVDPDPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240

Qy 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSPRDWQRLI 300
Db 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSPRDWQRLI 300

Qy 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360
Db 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360

Qy 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPPSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPPSQMLRTGNNFTSYTFEDVPF 420

Qy 421 HSSYAHQSGLDLMLNPLIDQYLYLSRNTTSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSYAHQSGLDLMLNPLIDQYLYLSRNTTSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480

Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540

Qy 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600
Db 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600

Qy 601 LFGMWQDRDVLQGPWQIWKIPHTDGHFHPSPLMGFGFLKHPPIQLIKNTVPANPSTT 660
Db 601 LFGMWQDRDVLQGPWQIWKIPHTDGHFHPSPLMGFGFLKHPPIQLIKNTVPANPSTT 660

Qy 661 FSAAKFASFIQYSTGQSVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFTVDNGVY 720
Db 661 FSAAKFASFIQYSTGQSVSEIWELOKENS KRNWPEIQYTSNKNKSVNVDFTVDNGVY 720

Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

RESULT 14

```
US-10-880-297-25
; Sequence 25, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAESTRI, NARENDRA
```

; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; CURRENT FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-25

Query Match 99.5%; Score 3975; DB 17; Length 735;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 731; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60

Qy 61 KGEVNEADAALAEHDKAYDRQLDSDGNPYLYKYNHADADEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAALAEHDKAYDRQLDSDGNPYLYKYNHADADEFQERLKEDTSFGNLRGAVFQ 120

Qy 121 AKKRVLEPLGLVEEVPKTAAGKRPVESHSPVPESSSGTGKAGQOPARKRLNFGOTGDAD 180
Db 121 AKKRVLEPLGLVEEVPKTAAGKRPVESHSPVPESSSGTGKAGQOPARKRLNFGOTGDAD 180

Qy 181 SVDPQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCDSTWMDRVI 240
Db 181 SVDPQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCDSTWMDRVI 240

Qy 241 TTSTRTWALPTYNHLYKQISSQASNDNHFGYSTPWGYDFNRFCHFSPRDWQRLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQASNDNHFGYSTPWGYDFNRFCHFSPRDWQRLI 300

Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVOVFTDSEYQLPVYLSAHOQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVOVFTDSEYQLPVYLSAHOQ 360

Qy 361 CLPPFPADVFPVQGYLTLLNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPFPADVFPVQGYLTLLNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTSYTFEDVPF 420

Qy 421 HSSVAHSQSLDRLMNPLIDQLYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNPLIDQLYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480

Qy 481 PCYQORVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFFPQSGVL 540
Db 481 PCYQORVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFFPQSGVL 540

Qy 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNEQAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNEQAATADVNTQGV 600

Qy 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILIKNTVPVNPSTT 660
Db 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILIKNTVPVNPSTT 660

Qy 661 FSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720

US-10-880-297-27
; Sequence 27, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESTRY, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; CURRENT FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-27

Query Match 99.5%; Score 3975; DB 17; Length 735;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60

Qy 61 KGEVNEADAALAEHDKAYDRQLDSDGNPYLYKYNHADADEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAALAEHDKAYDRQLDSDGNPYLYKYNHADADEFQERLKEDTSFGNLRGAVFQ 120

Qy 121 AKKRVLEPLGLVEEVPKTAAGKRPVESHSPVPESSSGTGKAGQOPARKRLNFGOTGDAD 180
Db 121 AKKRVLEPLGLVEEVPKTAAGKRPVESHSPVPESSSGTGKAGQOPARKRLNFGOTGDAD 180

Qy 181 SVDPQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCDSTWMDRVI 240
Db 181 SVDPQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCDSTWMDRVI 240

Qy 241 TTSTRTWALPTYNHLYKQISSQASNDNHFGYSTPWGYDFNRFCHFSPRDWQRLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQASNDNHFGYSTPWGYDFNRFCHFSPRDWQRLI 300

Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVOVFTDSEYQLPVYLSAHOQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVOVFTDSEYQLPVYLSAHOQ 360

Qy 361 CLPPFPADVFPVQGYLTLLNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPFPADVFPVQGYLTLLNGSQAVGRSSFYCLEYFPPSQMLRTGNNFTSYTFEDVPF 420

Qy 421 HSSVAHSQSLDRLMNPLIDQLYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNPLIDQLYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480

Qy 481 PCYQORVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFFPQSGVL 540
Db 481 PCYQORVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFFPQSGVL 540

Qy 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNEQAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNEQAATADVNTQGV 600

Qy 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILIKNTVPVNPSTT 660
Db 601 LPGMWQDRDYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILIKNTVPVNPSTT 660

Qy 661 FSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720

QY 721 SEPRIGTRYLRNL 735
Db |||||
721 SEPRIGTRYLRNL 735

Search completed: May 19, 2005, 13:46:05
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:24:55 ; Search time 42 Seconds
(without alignments)
1306.359 Million cell updates/sec

Title: US-10-038-972A-13
Perfect score: 3994
Sequence: 1 MAADGYLPDWLEDTLSEGI.....TNGVSEPRPIGTRYLTRNL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3994	100.0	735	US-09-321-589-1	Sequence 1, Appli
2	3994	100.0	735	US-10-293-478-1	Sequence 1, Appli
3	3402.5	85.2	736	US-09-807-802A-3	Sequence 3, Appli
4	3402.5	85.2	736	US-09-807-802A-13	Sequence 13, Appli
5	2764.5	69.2	599	US-09-807-802A-15	Sequence 15, Appli
6	2481.5	62.1	534	US-09-807-802A-17	Sequence 17, Appli
7	2397	60.0	734	US-09-532-594B-4	Sequence 4, Appli
8	1797	45.0	598	US-09-532-594B-16	Sequence 16, Appli
9	1669.5	41.8	544	US-09-532-594B-18	Sequence 18, Appli
10	1579.5	39.5	756	US-09-438-268-4	Sequence 4, Appli
11	486	12.2	543	US-08-856-841-22	Sequence 22, Appli
12	431	10.8	503	US-08-856-841-16	Sequence 16, Appli
13	431	10.8	501	US-08-856-841-18	Sequence 18, Appli
14	428	10.7	486	US-08-856-841-19	Sequence 19, Appli
15	327	8.2	415	US-08-856-841-20	Sequence 20, Appli
16	321	8.0	395	US-08-856-841-13	Sequence 13, Appli
17	312	7.8	264	US-08-856-841-14	Sequence 14, Appli
18	308	7.7	398	US-08-856-841-21	Sequence 21, Appli
19	282.5	7.1	387	US-08-856-841-17	Sequence 17, Appli
20	241.5	6.0	579	5223424-13	Patent No. 5223424
21	241.5	6.0	579	5223424-13	Patent No. 5223424
22	208	5.2	584	US-09-022-949-2	Sequence 2, Appli
23	138	3.5	227	US-08-856-841-15	Sequence 15, Appli
24	138	3.5	250	US-08-856-841-12	Sequence 12, Appli
25	136	3.4	210	US-08-856-841-9	Sequence 9, Appli
26	119	3.0	655	US-08-469-202-27	Sequence 27, Appli
27	119	3.0	655	US-08-484-434C-34	Sequence 34, Appli

ALIGNMENTS

RESULT 1

US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321.589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1
Query Match 100.0%; Score 3994; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAADGYLPDWLEDTLSEGIQWVKLKP	655	4	US-09-384-361-34	Sequence 34, Appli
			1095	4	US-09-107-532A-3855	Sequence 3855, Ap
Db	1	MAADGYLPDWLEDTLSEGIQWVKLKP	30	3	US-08-814-052-6	Sequence 6, Appli
			560	3	US-08-812-829-6	Sequence 6, Appli
Qy	61	KGEPVNEADAAALEHDKAYDQLDSD	31	3	US-09-373-157-4	Sequence 4, Appli
			560	3	US-09-373-157-4	Sequence 4, Appli
Db	61	KGEPVNEADAAALEHDKAYDQLDSD	32	4	US-09-949-016-8582	Sequence 8582, Ap
			847	4	US-09-386-962C-4	Sequence 4, Appli
Qy	121	AKKRVLEPLGLVEEPVKTPAGKGRV	33	4	US-09-134-001C-3865	Sequence 3865, Ap
			875	4	US-09-386-962C-4	Sequence 4, Appli
Db	121	AKKRVLEPLGLVEEPVKTPAGKGRV	34	4	US-09-710-279-1780	Sequence 1780, Ap
			1742	4	US-08-874-569B-21	Sequence 21, Appli
Qy	181	SVPDPQIPGQPPAASGLGNTMATG	35	4	US-09-955-518-21	Sequence 21, Appli
			501	4	US-09-248-796A-15119	Sequence 15119, A
Db	181	SVPDPQIPGQPPAASGLGNTMATG	40	4	US-09-538-092-1320	Sequence 1320, Ap
			983	4	US-09-248-796A-26692	Sequence 26692, A
Qy	241	TTSTRFTWALPYNNHLYKQISSQ	42	4	US-09-328-352-7017	Sequence 7017, Ap
			115.5	2.9	US-08-947-965-78	Sequence 78, Appli
Db	241	TTSTRFTWALPYNNHLYKQISSQ	43	3	US-08-469-202-28	Sequence 28, Appli
			624	3		
Qy	301	NNNGFRPKRLNFKLFTQVKEVTQ	44	111		
			655	1		
Db	301	NNNGFRPKRLNFKLFTQVKEVTQ	45			
Qy	361	CLPPPPADVFVMPQGYLTNNGS				
Db	361	CLPPPPADVFVMPQGYLTNNGS				

QY 421 HSSVAHSOSLDRLMNPILIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
DB 421 HSSVAHSOSLDRLMNPILIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
DB 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
QY 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
DB 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
QY 601 LPGVMQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPVAMPSTT 660
DB 601 LPGVMQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPVAMPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPETQYTSNYSKSVNVDFVTDTNGVY 720
DB 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPETQYTSNYSKSVNVDFVTDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735

RESULT 2

US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 100.0%; Score 3994; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYLPDWLEDTLSEGIROWKWKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIROWKWKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
QY 61 KGEVNEADAAALHDKAYDRQLDSGDNPLYLYNHADAERQERLKEDTSFGNIGRAVQ 120
DB 61 KGEVNEADAAALHDKAYDRQLDSGDNPLYLYNHADAERQERLKEDTSFGNIGRAVQ 120
QY 121 AKKRVLEPLGLVEPVTAPGKKRPVHSVPEDSSSGTGKAGQAPARKLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEPVTAPGKKRPVHSVPEDSSSGTGKAGQAPARKLNFGQTGDAD 180
QY 181 SVDPQPLQGPAPAPSGLGTNTMATGSGPAMDNNEGADGVNNSGNHCHDSTWMDRVI 240
DB 181 SVDPQPLQGPAPAPSGLGTNTMATGSGPAMDNNEGADGVNNSGNHCHDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYGYSTPMGYFDNFRHCHFSPRDWQRLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYGYSTPMGYFDNFRHCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360

DB 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
QY 361 CLPPPPADVFVMPQYGYLTLLNNGSAQVGRSSFYCLEYPPSQMLRTGNNTFSYTFEDVPF 420
DB 361 CLPPPPADVFVMPQYGYLTLLNNGSAQVGRSSFYCLEYPPSQMLRTGNNTFSYTFEDVPF 420
QY 421 HSSVAHSOSLDRLMNPILIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
DB 421 HSSVAHSOSLDRLMNPILIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
DB 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
QY 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
DB 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
QY 601 LPGVMQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPVAMPSTT 660
DB 601 LPGVMQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPVAMPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPETQYTSNYSKSVNVDFVTDTNGVY 720
DB 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPETQYTSNYSKSVNVDFVTDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735

RESULT 3

US-09-807-802A-3
; Sequence 3, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVEN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-3

Query Match 85.2%; Score 3402.5; DB 4; Length 736;
Best Local Similarity 83.3%; Pred. No. 2.2e-276;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
QY 1 MAADGYLPDWLEDTLSEGIROWKWKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIROWKWKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
QY 61 KGEVNEADAAALHDKAYDRQLDSGDNPLYLYNHADAERQERLKEDTSFGNIGRAVQ 120
DB 61 KGEVNEADAAALHDKAYDRQLDSGDNPLYLYNHADAERQERLKEDTSFGNIGRAVQ 120
QY 121 AKKRVLEPLGLVEPVTAPGKKRPVHSVPEDSSSGTGKAGQAPARKLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEPVTAPGKKRPVHSVPEDSSSGTGKAGQAPARKLNFGQTGDAD 180
QY 181 SVDPQPLQGPAPAPSGLGTNTMATGSGPAMDNNEGADGVNNSGNHCHDSTWMDRVI 240

Db	181	SVDPDQPLGEPPTATPAAGVPTTWASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRVI	240
Qy	241	TTSTRTWALPTYNNHLYKQISSQ-GASNDNHFYGYSTPWGYFDENRFCHFSPRDWORL	299
Db	241	TTSTRTWALPTYNNHLYKQISSASTCASNDNHFYGYSTPWGYFDENRFCHFSPRDWORL	300
Qy	300	INNNGFRPKRLNFKLFNQKQVTONDGTITIANNLTSVTQVFTDSEYQLPYVLGSAHQ	359
Db	301	INNNGFRPKRLNFKLFNQKQVTTNDGVTITIANNLTSVTQVFSDEYQLPYVLGSAHQ	360
Qy	360	GCLPPPPADVFMVPOXYGLTLNNGSOAVGRSSPYCLEYPPSQMLRTGNNFTFSYTFEDVP	419
Db	361	GCLPPPPADVFMIPQYGLTLNNGSAVGRSSPYCLEYPPSQMLRTGNNFTFSYTFEVP	420
Qy	420	FHSSYAHSSQLRLMPLIDQYLYLSRNTWPSGTTTQSRLOFSQAGASDIRDQSRNMLP	479
Db	421	FHSSYAHSSQLRLMPLIDQYLYLNRTCQSGSAQNKDLLFSRGSAGMSVQPKWLP	480
Qy	480	GPCYRQORVSKTSADNNNSYSGATKYKYLNGRDSLVPNGPAMASHKDDDEKFFPQSGV	539
Db	481	GPCYRQORVSKTKTDNNNSNFTWGASKYNLNGRESIINPGTAMASHKDDDEKFFPMGCV	540
Qy	540	LIFGKGSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRGNFQAATADVNTQG	599
Db	541	MIFGKSAGASNTALDNVMTDDEEIKATNPATERFGTVAVNFQSSSTDPTDGDVHAMG	600
Qy	600	VLFPMWQDRDVLQGPWIWAKIPHTDGHFHPSPLMGGFGLKHPPPOILLIKNTFPVANPST	659
Db	601	ALFGMWQDRDVLQGPWIWAKIPHTDGHFHPSPLMGGFGLKNPPPOILLIKNTFPVANPPA	660
Qy	660	TPSAAPKASPTQYSTQGVSVSEIWELOKENSKRWNPEIQYTSNYNKS VNVDFTVDVTNGV	719
Db	661	EPSATKFASPTQYSTQGVSVSEIWELOKENSKRWNPEVQYTSNYAKSANVDFTDVNNGL	720
Qy	720	YSEPRPIGTGTRYLRNL	735
Db	721	YTEPRPIGTGTRYLRNL	736
RESULT 4			
US-09-807-802A-13			
; Sequence 13, Application US/09807802A			
; Patent No. 6759237			
; GENERAL INFORMATION:			
; APPLICANT: Wilson, James M.			
; APPLICANT: Xiao, Weidong			
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,			
; TITLE OF INVENTION: Vectors and Host Cells Containing Same			
; FILE REFERENCE: GNVPN.031USA			
; CURRENT APPLICATION NUMBER: US/09/807,802A			
; CURRENT FILING DATE: 2002-02-21			
; PRIOR APPLICATION NUMBER: US 60/107,114			
; PRIOR FILING DATE: 1998-11-05			
; PRIOR APPLICATION NUMBER: PCT/US99/25694			
; PRIOR FILING DATE: 1999-11-02			
; NUMBER OF SEQ ID NOS: 20			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 13			
; LENGTH: 736			
; TYPE: PRM			
; ORGANISM: AAV-1			
US-09-807-802A-13			
Query Match 85.2%; Score 3402.5; DB 4; Length 736;			
Best Local Similarity 83.3%; Pred.No.2.2e-276;			
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;			
Qy	1	MAADGYLPDWLEDTLSEGIROWKLPKGP PPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWLENLSEGIREWDLKPGAPKPKANQKQDDRGRLVLPGYKYLGPFGNGLD	60
Qy	61	KGEFPVNEADAAALHDKAYDRQLDSDGNPPLYKTNHADAEFQERLKDFTSFGGNLGRAVPQ	120

Db 61 KGPVNAADAALAEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSPGGNLGRAVFQ 120
 Qy 121 AKKRVLEPLGLVBEPPVKTAPEKKRPVEHSSPVEDSSGTCAGAQOQPAKRLNFGQTGDAD 180
 Db 121 AKKRVLEPLGLVBEAGAKTAPEKKRPVEQSPQEPDSSGIGKTCQOQPAKRLNFGQTGDSE 180
 Qy 181 SVDPDQPLGPPAAPSGLGNTWTATSGAPMADNEGADGVGNSGNGHCHDSTWGMGRVI 240
 Db 181 SVDPDQPLGPPAPTAAVGGTTTWSAGSGAPMADNEGADGVGNSAGNGHCHDSTWLGDRVI 240
 Qy 241 TTSRTWTALPTYNNHLYKQISSLSS-GASNDNHVFGYSTPMGYDFDNFPHCHFSPRDQRL 299
 Db 241 TTSRTWTALPTYNNHLYKQISSLASSTGASNDNHVFGYSTPMGYDFDNFPHCHFSPRDQRL 300
 Qy 300 INNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 359
 Db 301 INNNWGFRPKRLNFKLFNIQVKEVTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQ 360
 Qy 360 GCLPPPPADVFMVPOGYLYTLNNGSQAAGRSSFYCLEYFPFSOMLRTGNNTFTSFYTFEDVP 419
 Db 361 GCLPPPPADVFMIPQGYLYTLNNGSQAAGRSSFYCLEYFPFSOMLRTGNNTFTSFYTFEVP 420
 Qy 420 FHSSYAHSSQLDRMLNPLIDQYLYLVRWTTPSGTTTQSLRQSOAGASDIRDQSRNWL 479
 Db 421 FHSSYAHSSQLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKNWL 480
 Qy 480 GPCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGV 539
 Db 481 GPCYRQORVSKTKTDNNNSFTWTGASKYNLNGRESIIINEGTAMASHKDDDEKFFPMSGV 540
 Qy 540 LIFGKGSEKTNVDIEKVMITDDEEIRTNTPVATEQYGVSVTLQNRGNOQAATADVNTQG 599
 Db 541 MIFGKSAGASNTALDNVMTDDEEIKATNPVATERPGTVAVNFQSSSTDPATGCDVHAMG 600
 Qy 600 VLPGMWQDRDVLQGPFIWAKIPIHTDGHFHPSPLMGFGFLKHPPQQLIKNTPVPANPST 659
 Db 601 ALPGMWQDRDVLQGPFIWAKIPIHTDGHFHPSPLMGFGFLKHPPQQLIKNTPVPANPPA 660
 Qy 660 TFSAAKFPASPIQYSTGVQSVSEIWELOKENSXRNWPEIQYTSNYNKSXNVVDFTVDTNGV 719
 Db 661 EFSATKFPASPIQYSTGVQSVSEIWELOKENSXRNWPEVQYTSNYAKSANVDFTVDNGL 720
 Qy 720 YSEPRPIGTRYLTRLN 735
 Db 721 YTEPRPIGTRYLTRPL 736
 RESULT 5
 US-09-807-802A-15
 ; Sequence 15, Application US/09807802A
 ; Patent No. 6759237
 ; GENERAL INFORMATION:
 ; APPLICANT: Wilson, James M.
 ; APPLICANT: Xiao, Weidong
 ; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences
 ; TITLE OF INVENTION: Vectors and Host Cells Containing Same
 ; FILE REFERENCE: GNVN 031USA
 ; CURRENT APPLICATION NUMBER: US/09/807, 802A
 ; CURRENT FILING DATE: 2002-02-21
 ; PRIOR APPLICATION NUMBER: US 60/107,114
 ; PRIOR FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/25694
 ; PRIOR FILING DATE: 1999-11-02
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 599
 ; TYPE: PRT
 ; ORGANISM: AAV-1
 US-09-807-802A-15

Query Match	69.2%;	Score 2764.5;	DB 4;	Length 599;
Best Local Similarity	82.8%;	Pred. No. 5.1e-223;		

Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
QY 138 TAPGKRPVEHSVPPEPSSSGTGKAGQAPARKLNFGQGDADSVDPDPOPLGQPPAASG 197
Db 1 TAPGKRPVQSGPQEPDSSSGIGKTGQPAKRLNFGQGDSESVDPDPOPLGQPPA 60
QY 198 LGTMTATGSGAPMADNNEGADGVGNSGNWCHDSTWMDRVTITSTRWALPTYNHLY 257
Db 61 VGPTTASGGAPMADNNEGADGVGNSGNWCHDSTWMDRVTITSTRWALPTYNHLY 120
QY 258 KQISSOS-GASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGFPRKLNKLP 316
Db 121 KQISSASTGASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGFPRKLNKLP 180
QY 317 NIOQKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVLSAHQGLPPFPADVPMVPOYG 376
Db 181 NIOQKEVTNDGVTITIANNLTSVQVFTDSEYQLPVLSAHQGLPPFPADVPMVPOYG 240
QY 377 YLTNNGSOAVGRSSFYCLEYFPSSQMLRTGNFTSYTFEDVPHSSYAHQSQSLDRLMNP 436
Db 241 YLTNNGSOAVGRSSFYCLEYFPSSQMLRTGNFTSYTFEDVPHSSYAHQSQSLDRLMNP 300
QY 437 LIDQLYLNRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 496
Db 301 LIDQLYLNRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 360
QY 497 NSEYSWTGATKYHLNDRSLVNPMPAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEK 556
Db 361 NSNFTWQASYNLNGRESIINPGTAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEK 514
QY 557 VMITDEEIRTNVATEQYSGVSTNLQRNQRAATADVNTQGLVPGMWQDRDVLQGP 616
Db 421 VMITDEEIKATNPVATERGTVAVNFQSSSTDPATGVDHMGALPGMWQDRDVLQGP 480
QY 617 IWAKIPHTDGHFHPSPLMGFGFKHPPPQILIKNTVPANPSTTFSAAKFAFSTQYSG 676
Db 481 IWAKIPHTDGHFHPSPLMGFGFKHPPPQILIKNTVPANPSTTFSAAKFAFSTQYSG 540
QY 677 QVSVEIWELOKSKRNWPEIQYTSNYSKNSVNVDTVDNTNGVSEPRPIGTRYLTRNL 735
Db 541 QVSVEIWELOKSKRNWPEIQYTSNYSKNSVNVDTVDNTNGVSEPRPIGTRYLTRNL 599

RESULT 6
US-09-807-802A-17
; Sequence 17, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807.802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 534
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-17

Query Match
Best Local Similarity 62.1%; Score 2481.5; DB 4; Length 534;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;
QY 203 MATGSGAPMADNNEGADGVGNSGNWCHDSTWMDRVTITSTRWALPTYNHLYKQISS 262

Db 1 MASGGGAPMADNNEGADGVGNSGNWCHDSTWMDRVTITSTRWALPTYNHLYKQISS 60
QY 263 QS-GASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGFPRKLNKLPFNIOVK 321
Db 61 ASTGASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGFPRKLNKLPFNIOVK 120
QY 322 EYVNDGTTTIANNLTSVQVFTDSEYQLPVLSAHQGLPPFPADVPMVPOYGILTN 381
Db 121 EYVNDGVTITIANNLTSVQVFTDSEYQLPVLSAHQGLPPFPADVPMVPOYGILTN 180
QY 382 NGSQAVGRSSFYCLEYFPSSQMLRTGNFTSYTFEDVPHSSYAHQSQSLDRLMNPIDQY 441
Db 181 NGSQAVGRSSFYCLEYFPSSQMLRTGNFTSYTFEDVPHSSYAHQSQSLDRLMNPIDQY 240
QY 442 LYLSRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNSEYS 501
Db 241 LYLSRNTQNSGSAQNKDLLFSRSGPAGMSVQPKNWLPGPCYRQQRVSKTIDNNNSFT 300
QY 502 WTGATKYHLNDRSLVNPMPAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 561
Db 301 WTGATKYHLNDRSLVNPMPAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 360
QY 562 EBEIRTNVATEQYSGVSTNLQRNQRAATADVNTQGLVPGMWQDRDVLQGPIMAKI 621
Db 361 EBEIRKATNPVATERGTVAVNFQSSSTDPATGVDHMGALPGMWQDRDVLQGPIMAKI 420
QY 622 PHTDGHFHPSPLMGFGFKHPPPQILIKNTVPANPSTTFSAAKFAFSTQYSGQVSVE 681
Db 421 PHTDGHFHPSPLMGFGFKHPPPQILIKNTVPANPSTTFSAAKFAFSTQYSGQVSVE 480
QY 682 IEWELQKSKRNWPEIQYTSNYSKNSVNVDTVDNTNGVSEPRPIGTRYLTRNL 735
Db 481 IEWELQKSKRNWPEIQYTSNYSKNSVNVDTVDNTNGVSEPRPIGTRYLTRNL 534

RESULT 7
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532.594B
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VPI
US-09-532-594B-4

Query Match
Best Local Similarity 60.0%; Score 2397; DB 4; Length 734;
Matches 457; Conservative 90; Mismatches 167; Indels 36; Gaps 13;
QY 4 DGYLPDWLEDTLSEGIROWKLPKPPPPKPAERHKDDSKGLVLPGYKYIGPFGNGLDKGE 63
Db 3 DGYLPDWLEDTLSEGIROWKLPKPPPPKPAERHKDDSKGLVLPGYKYIGPFGNGLDKGE 62
QY 64 PVNEADAALHDKAYDQDLSGDNPYLYKNHDAEFQERLKEDTSGNGLGRAVFOAKK 123
Db 63 PVNEADAALHDKAYDQDLSGDNPYLYKNHDAEFQERLKEDTSGNGLGRAVFOAKK 122

```
QY 124 RVLEPLGLVEBPVKAPGKRVPVEHSPVPEPDSSTGKAGQOPARKRLNF--GOTGDADSV 182
DB 123 RVLEPLGLVEQAGETAPGKRPLIESPQPDSTGIGKKGQOPAKKLVFEDETGAGD-- 180
QY 183 PDPQPLGQPPAAPAGSL--GTTNTATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
DB 181 -----GPEGSTSGMADSEKRAAAGAAVEGGQADGVGNAGSDWHCDSTWSEGHVT 234
QY 241 TTSTRTWALPTYNHLYKOISSQSGASNDNHFGYSTPMGYFDNRFHCFHSPRDMQRLI 300
DB 235 TTSTRTWALPTYNHLYKRLGE---SLQSNNTYNGFSTPMGYFDNRFHCFHSPRDMQRLI 291
QY 301 NNNWGRPKRLNFKLFIQKVEVTQNDGTTTIANNLITSTVQVFTDSEYQLPYVLGSAHQG 360
DB 292 NNNWGRPKRAMPVKLFIQKVEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQSG 351
QY 361 CLPPPPADVFMVPOQGY---LTLNNGSAQVGRSSFYCLFYFSPQMLRTGNNTFTFSTRED 417
DB 352 SLPPFPNDVFMVPOQGYCGLVNTGNTSQOQTDNRNPFYCLFYFSPQMLRTGNNTFTFSTRED 411
QY 418 VPFHSSYAHQSGLDRMLNPLIDQYLYLSRTNTPSGTTTQ---SRLQFSQAGASDIRDQS 474
DB 412 VPFHSMYAHQSGLDRMLNPLIDQYLGWLGQSTTT--GTTLNAGTATTNFTKLRTNFSNPK 469
QY 475 RNWLPGPCVRQORVSKTSADNNSEYSWTGA---TKYH---LNGRDSLVPNPGPAMASHK 527
DB 470 KMWLPGPSIKQOGFSKTA--NQNYKIPATGSDSLIKYETHSLDGRWSALTGPPMATAG 527
QY 528 DEEEKFFPQSGVLIFG--KQSEKTNVDIEKVMITDEEIRTNPNVATEQYGSVSTNLR 585
DB 528 PADSKF--SNSQLIFAGPKQNGTATVP--GTLIFTSEELAATNATDTDMWGNLPGDQGS 584
QY 586 GNRQAATAADVNTQGVLPQVMWQDRDVIYQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQ 645
DB 585 NSNLPTVDRLTALGAVPGVMWQNRDIYQGPWAKIPHTDGHFHPSPLLGGFGLKHPPPQ 644
QY 646 ILIKNTPVPANPSTTFSAAKFASFTQYSTGVSVSVEIWELOKENSKRWNPEIQVTSNNY 705
DB 645 IPKNTVPANPATTFSSFPVNSFTQYSTGVSVQIDWEIQKRSKRWNPEVQVTSNYG 704
QY 706 KSVNVDFVTDTNGVYSEPRPICTRYLTRNL 735
DB 705 QNSLWAPDAAGKYTEPRAIGTRYLTRHL 734
```

RESULT 8

```
US-09-532-594B-16
; Sequence 16, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; * CURRENT FILING DATE: 2000-03-22
; * NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP2
US-09-532-594B-16
```

Query Match 45.0%; Score 1797; DB 4; Length 598;
Best Local Similarity 56.7%; Pred. No. 6.9e-142;

```
Matches 349; Conservative 76; Mismatches 155; Indels 36; Gaps 13;
QY 138 TAPGKRVPVEHSPVPEPDSSTGKAGQOPARKRLNF--GOTGDADSVDPQPLGQPPAADS 196
DB 1 TAPGKRPLIESPQPDSTGIGKKGQOPAKKLVFEDETGAGD-----GPEGSTS 52
QY 197 GL--GTTNTATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVIITSTRTWALPTYN 254
DB 53 GMSDDSEMRAAAGAAVEGGQADGVGNAGSDWHCDSTWSEGHVTTTSTRTWALPTYN 112
QY 255 HLYKOISSQSGASNDNHFGYSTPMGYFDNRFHCFHSPRDMQRLINNNWGRPKRLNPK 314
DB 113 HLYKRLGE---SLQSNNTYNGFSTPMGYFDNRFHCFHSPRDMQRLINNNWGRPKRMPVK 169
QY 315 LFIQKVEVTQNDGTTTIANNLITSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMVPO 374
DB 170 LFIQKVEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQSGSLPPFPNDVFMVPO 229
QY 375 YGY---LTLNNGSAQVGRSSFYCLFYFSPQMLRTGNNTFTFSTREDVFPFHSSYAHQS 431
DB 230 YGYCGLVNTGNTSQOQTDNRNPFYCLFYFSPQMLRTGNNTFTFSTREDVFPFHSSYAHQS 289
QY 432 RLMPNPLIDQYLYLSRTNTPSGTTTQ---SRLQFSQAGASDIRDQSRNWLPGPCYRQORV 488
DB 230 RLMPNPLIDQYLGWLGQSTTT--GTTLNAGTATTNFTKLRTNFSNPKMWLPGPSIKQOGF 347
QY 489 SKTSADNNSEYSWTGA---TKYH---LNGRDSLVPNPGPAMASHKDDEEKFFPQSGVLI 541
DB 348 SKTA--NQNYKIPATGSDSLIKYETHSLDGRWSALTGPPMATAGPADSKF--SNSQLI 403
QY 542 FG--KQSEKTNVDIEKVMITDEEIRTNPNVATEQYGSVSTNLRQNRQAATAADVNTQ 599
DB 404 FAGPKQNGTATVP--GTLIFTSEELAATNATDTDMWGNLPGDQGSNLTPTVDRLTALG 462
QY 600 VLPQVMWQDRDVIYQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPST 659
DB 463 AVPGVMWQNRDIYQGPWAKIPHTDGHFHPSPLLGGFGLKHPPPQIFIKNTPVPANPAT 522
QY 660 TFSAAKFPASFTQYSTGVSVSVEIWELOKENSKRWNPEIQVTSNNYKSVNVDFVTDTNGV 719
DB 523 TFSSTPVPNSFTQYSTGVSVQIDWEIQKRSKRWNPEVQVTSNYGQNSLWAPDAAGK 582
QY 720 YSEPRPICTRYLTRNL 735
DB 583 YTEPRAIGTRYLTRHL 598
```

RESULT 9

```
US-09-532-594B-18
; Sequence 18, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; * CURRENT FILING DATE: 2000-03-22
; * NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3
US-09-532-594B-18
```


COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTION)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 12.2%; Score 486; DB 3; Length 543;
Best Local Similarity 27.4%; Pred. No. 5,1e-32;
Matches 148; Conservative 80; Mismatches 221; Indels 92; Gaps 17;
Qy 202 TWATGS-----GAPMADNNEGADGVNSS--GNWHCDSTWMDGRVITTRTWTALPTYNH 255
Db 2 TWITNSLMTSVNSAEASTGAGGGGNSVKMSWSEGAFTSANSVCTCTFSRQFLIPYDPEH 61
Qy 256 LYKQISSQSGASND-----NHVFGYSTPMGYDFENRPFCHFSFPRDWORLINNW 304
Db 62 HYKVFSPAASCHNAGKEAKVCTTISPMGYSTPMRYLDFNALNLFPSLRFQHLIENYG 121

RESULT 12

US-08-856-841-16
; Sequence 16, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:

DB	491	PL	492
TELEPHONE: (212) 697-3355			
TELEFAX: (212) 557-5635			
TELEX: NONE			
INFORMATION FOR SEQ ID NO: 18:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 501			
TYPE: AMINO ACID			
TOPOLOGY: LINEAR			
MOLECULE TYPE:			
DESCRIPTION: PEPTIDE			
HYPOTHETICAL: N/A			
ANTI-SENSE: N/A			
FRAGMENT TYPE: INTERNAL			
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE			
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIONUM)			
POSITION IN GENOME: N/A			
FEATURE:			
NAME/KEY: N/A			
LOCATION: N/A			
IDENTIFICATION METHOD: amino acid analysis and			
IDENTIFICATION METHOD: mass spectrometry			
OTHER INFORMATION:			
PUBLICATION INFORMATION:			
AUTHORS: COSSART, Y.E.			
AUTHORS: FIELD, A.M.			
AUTHORS: CANT, B.			
AUTHORS: WIDDOWS, D.			
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA			
JOURNAL: LANCET			
VOLUME: 1			
ISSUE:			
PAGES: 72 - 73			
DATE: 1975			
DOCUMENT NUMBER:			
FILING DATE:			
PUBLICATION DATE:			
RELEVANT RESIDUES IN SEQ ID NO: 18:			
US-08-856-841-18			
Query Match	10.88;	Score 431;	DB 3; Length 501;
Best Local Similarity	27.58;	Pred. No. 1.8e-27;	
Matches 116; Conservative	50;	Mismatches 180;	Indels 76; Gaps 7;
Qy	38	HKDDSRGLVLP	PGYKLGPFNGDKGPEVNEADAAALEHDKAYDRQLDSDGNPYLKNHAD 97
Db	125	HKPGQVSVQVLP	PGTYVCGFNGELQNGPQSAVDSNARIHDFRYSQLAKLGINPTHTWTVAD 184
Qy	98	AEQERLEKEDTS	FGNIGRAVFAQKRVLEPLGLVIEEPVKTAPKRPVHSPVPEPDSS 157
Db	185	EELLKNIKNETG	FQAVVKDYF-----TLGMAAPVAHF----- 218
Qy	158	GTGKAGQAPARK	LNFGQTGDADVDPDPLGQPAPSGLGTNTWATGSGAPMADNNEG 217
Pb	219	-----	QGSLPEVPAYNAASEKYPMSMTSVNSAEASTGA-----G 250
Qy	218	ADGVGNSSGNW	CHDSTWMDRVLTSTRTWALPTYNNHLYKQISSQSGASND----- 269
Db	251	GGGNSVKMSW	SEGAFTSANSVTCTSRQQLIPVDPEHHYKVFSPAASSCHNAGKEAKV 310
Qy	270	---NHVFGYST	PMGYFDNRFHCFSPRDQRLINNHWGFRPKRLNFKLFNIQVKEVT-Q 325
Db	311	CTISPIMGYST	PMRYLDNALNLFSPLEFQHLIENYGSTAPDALVTIIEIAVKDVTDK 370
Qy	326	NDGTTTIANLT	STVQVFTSEYQLPVYLSAHQGCCLPPFPADVFMVQVGYTLNN-GS 384
Db	371	TGGGVQVDTST	TGRLCLMLVDHEYKPYVYLGQGDQTLAPELPIWYFPQVAYLTGVDNT 430
Qy	385	QAVG-----	RSFFCYCLEYFSPSOMLRGNFTFSTYTFEDVPFHSSYHQSQDLNWN 435
Db	431	QGISGDSK	KLASEEAFVYVLEHSSFQILGTGGTASMSYKFPFPPPENLECGSCHFYEMYN 490
Qy	436	PL	437

Qy 436 PL 437

RESULT 15

Db 184 NTRISL-RPGVSPYHHWDTDKYVTGINAISHGQTTYGNAEDKEYQQGVGRFPNEKEQL 242
QY 558 MITDEEEIRTTNP-VATEOYGSVSTNLQRGNQAAATADVNTQGVLPGMVWQDRDVYLOGP 616
Db 243 KOLOGLNWHITYFPNKCTQY-----TDQIERPLMVGSVWNRRAHYESQ 286
QY 617 IWAKIPHTDGHFHS-PLMGFGGLKHPPPOIILKNTPVPANPSTTFSAAKFAFIT--QY 673
Db 287 LWSKIPNLDDSFKTOFAALGGWGLHQPPQIFLKILP-----ESGPIGGIKSMGITTLVQY 342
QY 674 STGOVSVEIEWEL-OKENSKRWNP 697
Db 343 AVGIMTVTMTFKLGRKATGRWN PQ 367

Search completed: May 19, 2005, 13:34:37
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:15:05 ; Search time 165 Seconds
(without alignments)
1722.842 Million cell updates/sec

Title: US-10-038-972A-13

Perfect score: 3994

Sequence: 1 MAADGYLPDWLEDTLSEGR.....TNGVSEPRPIGTRYLTRNL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3994	100.0	735	4	AAG65792 Adeno-ss
2	3994	100.0	735	4	AAB59844 AAV2 caps
3	3994	100.0	735	4	AAM51508 Adeno-ss
4	3994	100.0	735	5	AAY98974 Adeno-ss
5	3994	100.0	735	6	ABR80231 AAV2 vpl
6	3994	100.0	735	6	ABR82436 Adeno-ss
7	3994	100.0	735	7	ABR62761 Adeno-ss
8	3994	100.0	735	7	ADP76571 Adeno-ss
9	3967.5	99.3	734	4	AAB50326 Adeno-ss
10	3623	90.7	735	7	ADP76568 Adeno-ss
11	3622	90.7	735	7	ADP76570 Adeno-ss
12	3617	90.6	735	7	ADP76569 Adeno-ss
13	3615	90.5	735	7	ADP76567 Adeno-ss
14	3557.5	89.1	736	4	AAB59846 AAV3B cap
15	3532.5	88.4	736	4	AAB59845 AAV3A cap
16	3532.5	88.4	736	6	ABR80233 AAV3 vpl
17	3532.5	88.4	736	7	ABR62763 Adeno-ss
18	3532.5	88.4	736	7	ADP76572 Adeno-ss
19	3421	85.7	731	7	ADP76589 Adeno-ss
20	3416.5	85.5	738	7	ADP76592 Adeno-ss
21	3409.5	85.4	736	4	AAB59847 AAV6 caps
22	3409.5	85.4	736	7	ADP76566 Adeno-ss
23	3405.5	85.3	738	7	ADP76594 Adeno-ss
24	3405.5	85.3	738	7	ADP76594 Adeno-ss
25	3404.5	85.2	738	7	ADP76580 Adeno-ss

26	3404.5	85.2	738	7	ADP76586	Ade76586 Adeno-ss
27	3404.5	85.2	738	7	ADP76582	Ade76582 Adeno-ss
28	3402.5	85.2	736	3	AAV71167	AAV71167 Adeno-ss
29	3402.5	85.2	736	6	ABR80232	ABR80232 AAV1 vpl
30	3402.5	85.2	736	7	ABR62762	ABR62762 Adeno-ss
31	3402.5	85.2	736	7	ADP76565	Ade76565 Adeno-ss
32	3401	85.2	731	7	ADP76590	Ade76590 Adeno-ss
33	3398.5	85.1	738	7	ADP76593	Ade76593 Adeno-ss
34	3397.5	85.1	738	7	ADP76581	Ade76581 Adeno-ss
35	3396	85.0	733	7	ADP76591	Ade76591 Adeno-ss
36	3392.5	84.9	736	7	ADP76598	Ade76598 Adeno-ss
37	3392.5	84.9	738	7	ADP76583	Ade76583 Adeno-ss
38	3391.5	84.9	736	6	ABR80229	ABR80229 AAV9 cap
39	3391.5	84.9	736	7	ADP76584	Ade76584 Adeno-ss
40	3391.5	84.9	736	7	ADP76601	Ade76601 Adeno-ss
41	3391.5	84.9	738	7	ADP76585	Ade76585 Adeno-ss
42	3386.5	84.8	738	7	ADP76584	Ade76584 Adeno-ss
43	3383.5	84.7	736	7	ADP76597	Ade76597 Adeno-ss
44	3381	84.7	733	7	ADP76588	Ade76588 Adeno-ss
45	3379.5	84.6	736	7	ADP76599	Ade76599 Adeno-ss

ALIGNMENTS

RESULT 1

AAG65792

ID AAG65792 standard; protein; 735 AA.

XX AC AAG65792;

XX DT 11-FEB-2002 (first entry)

XX DE Adeno-associated virus 2 (AAV-2) major coat protein VPI.

XX KW Recombinant viral vector; RVV; capsid; parvovirus; transgene; cytostatic; inverted terminal repeat; neuroprotective; antianemic; ITR; antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV; major coat protein; AAV-2; VPI.

XX OS Adeno-associated virus 2.

XX PN WO200168888-A2.

XX PD 20-SEP-2001.

XX PF 13-MAR-2001; 2001WO-US007927.

XX PR 14-MAR-2000; 2000US-0189110P.

XX PA (NEUR-) NEUROLOGIX INC.

XX PI Xiao W, During MJ;

XX DR WPI; 2001-596912/67.

XX DR N-PSDB; AAI66974.

XX PT Recombinant viral vector useful in improving gene therapy in a subject, and for increasing efficiency of entry into a cell, comprises a chimeric capsid having one non-native amino acid sequence and a desired transgene.

XX PS Disclosure; Page 50; 53pp; English.

XX CC The invention provides a recombinant viral vector (RVV) comprising a chimeric capsid (I) having at least one non-native amino acid sequence, derived from a capsid protein domain of parvovirus (II), a virus (III), or their combination, and a transgene flanked 5' and 3' by inverted terminal repeat (ITR) sequences, derived from (II), (III), or their combination. The RVV is useful for improving gene therapy in a subject with a disorder, and for increasing the efficiency of entry into a cell, which involves producing (I) encapsulating a viral vector, and contacting a cell with RVV having (I) such that (I) binds to an attachment site on the cell surface and permits the vector to enter the cell efficiently. A

pharmaceutical composition comprising RVV with (I) containing a transgene sequence associated with a disease or a disorder such that expression of the transgene would result in amelioration of the disease or disorder such as inherited neurological and metabolic diseases e.g. lysosomal storage disease, Leach-Nyhan syndrome, amyloid polynuropathy, Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood e.g. sickle cell anemia, clotting disorders and thalassemias, cystic fibrosis, diabetes, diseases associated with hormone deficiencies, retinoblastoma and various types of neoplastic cells which include tumours especially central nervous system tumours, neoplasms, carcinomas, sarcomas, leukemias and lymphoma. The present sequence represents the adeno-associated virus 2 (AAV-2) major coat protein VP1. AAV-2 sequences are used in the construction of a chimeric vector

Query Match 100.0%; Score 3994; DB 4; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60
 DB 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60

QY 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKDTSFGNIGRAVFQ 120
 DB 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKDTSFGNIGRAVFQ 120

QY 121 AKKRVLEPLGLVEEPVKTAPGKKRVEHSPVPEPSSSGTGKAGQOPARKRLNFGQTGDAD 180
 DB 121 AKKRVLEPLGLVEEPVKTAPGKKRVEHSPVPEPSSSGTGKAGQOPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLQOPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240
 DB 181 SVDPDQPLQOPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240

QY 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHYFGYSTPWGTFDNRPHCHFSPRDWRLI 300
 DB 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHYFGYSTPWGTFDNRPHCHFSPRDWRLI 300

QY 301 NNNWFRPKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
 DB 301 NNNWFRPKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360

QY 361 CLPPPADVFMVQGYLTLANGQAVGRSSFCYCLEYFPSSQMLRTGNNTFSYTFEDVPF 420
 DB 361 CLPPPADVFMVQGYLTLANGQAVGRSSFCYCLEYFPSSQMLRTGNNTFSYTFEDVPF 420

QY 421 HSSVAHSQSLDRLMNPILDQLYLSTNTPTSGTTTOSRLOFSQAGASDIRDOSRWLPG 480
 DB 421 HSSVAHSQSLDRLMNPILDQLYLSTNTPTSGTTTOSRLOFSQAGASDIRDOSRWLPG 480

QY 481 PCYQQRVSKTSADNNSEYSWTQATKYHLNGRDSLVPNPGPAMASHKDDBEKFFPQSGVL 540
 DB 481 PCYQQRVSKTSADNNSEYSWTQATKYHLNGRDSLVPNPGPAMASHKDDBEKFFPQSGVL 540

QY 541 IFGQSGSEKTNVDIEKWIITDEEIRTNVATEQYGSVSTNLRGNRQAATADVNTQGV 600
 DB 541 IFGQSGSEKTNVDIEKWIITDEEIRTNVATEQYGSVSTNLRGNRQAATADVNTQGV 600

QY 601 LPMVMDQDRDYLQGPQIWKAKIPTHGDFHPSPLMGGFGLKHPPPQILLIKNTVPANPSIT 660
 DB 601 LPMVMDQDRDYLQGPQIWKAKIPTHGDFHPSPLMGGFGLKHPPPQILLIKNTVPANPSIT 660

QY 661 FSAKFAFSTQYSTGVSEIWELOKENSKNWNPETQVTSYNNKSVNVDFVDTNGVY 720
 DB 661 FSAKFAFSTQYSTGVSEIWELOKENSKNWNPETQVTSYNNKSVNVDFVDTNGVY 720

QY 721 SEPRPIGTRILTRNL 735
 DB 721 SEPRPIGTRILTRNL 735

RESULT 2
 AAB59844
 ID AAB59844 standard; protein; 735 AA.
 AC AAB59844;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE AAV2 capsid protein VP1.
 XX
 KW AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
 KW atherosclerosis; sickle cell anaemia; thalassemia;
 KW blood clotting disorder; diabetes; capsid protein VP1.
 XX
 OS Adeno associated virus.
 XX
 PN US6156303-A.
 XX
 PD 05-DEC-2000.
 XX
 PF 11-JUN-1997; 97US-00873168.
 XX
 PR 11-JUN-1997; 97US-00873168.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Russell DW, Rutledge EA;
 FI
 DR WPI; 2001-060164/07.
 XX
 CC Adeno-associated virus serotype 6 and viral vector derived from it for
 CC gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
 CC syndrome, sickle cell anemia, thalassemia and diabetes.
 CC
 PS Claim 7; Fig 2; 50pp; English.
 XX
 CC The present invention relates to adeno-associated virus serotypes. The
 CC present sequence is capsid protein VP1 of one such serotype (AAV2). AAV2
 CC can be used to construct AAV viral vectors for use in gene therapy for a
 CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
 CC sickle cell anaemia, thalassemia, blood clotting disorders and diabetes.
 CC The AAV viral vectors have increased transduction efficiency of a
 CC particular host cell as the AAV virion containing the AAV vector genome
 CC can be modified to express a capsid protein of an AAV serotype that
 CC transduces the selected host cell
 XX
 SQ Sequence 735 AA;

Query Match 100.0%; Score 3994; DB 4; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60
 DB 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60

QY 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKDTSFGNIGRAVFQ 120
 DB 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKDTSFGNIGRAVFQ 120

QY 121 AKKRVLEPLGLVEEPVKTAPGKKRVEHSPVPEPSSSGTGKAGQOPARKRLNFGQTGDAD 180
 DB 121 AKKRVLEPLGLVEEPVKTAPGKKRVEHSPVPEPSSSGTGKAGQOPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLQOPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240
 DB 181 SVDPDQPLQOPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNHCHDSTWMDRVI 240

QY 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHYFGYSTPWGTFDNRPHCHFSPRDWRLI 300
 DB 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHYFGYSTPWGTFDNRPHCHFSPRDWRLI 300

QY 301 NNNWFRPKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360

PN WO200253703-A2.
 XX 11-JUL-2002.
 XX 04-JAN-2002; 2002WO-US000152.
 XX 05-JAN-2001; 2001US-0260124P.
 XX (CHIL-) CHILDRENS HOSPITAL INC.
 XX Bartlett JS;
 XX WPI; 2002-583608/62.
 XX N-P5DB; ABK89694.
 XX
 XX New adeno-associated virus vector comprises a biotinylated capsid or
 XX capsid protein with an amino acid insertion in the vp1 capsid, useful as
 XX a vaccine or for transferring a therapeutic peptide to a cancer cell.
 XX
 XX Claim 1; Page 49-51; 57pp; English.
 XX
 XX The invention relates to an adeno-associated virus (AAV) vector (I)
 XX comprising a biotinylated capsid or capsid protein (II) with an amino
 XX acid insertion following the capsid amino acid at position 139, 161, 588
 XX or 657 in the vp1 capsid. The AAV vector comprises a capsid protein
 XX containing one or more amino acid insertions that ablate the ability of
 XX the vector to bind heparin-sulphate proteoglycan and allow the vector to
 XX use a cellular receptor not used by wild type AAV. Modified (II) are
 XX useful as vaccines to elicit immune responses to amino acids, where the
 XX response can be protective and/or therapeutic. (I) may be used to
 XX transfer a therapeutic peptide to a cancer cell, particularly to an
 XX ovarian cancer cell. The present sequence represents the adeno-associated
 XX virus 2 (AAV2) vector, vp1 capsid protein used in to make modified AAV2
 XX vectors
 XX
 XX SQ Sequence 735 AA;
 Query Match 100.0%; Score 3994; DB 5; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAADGYLPDMLBTLSEGIQWMLKPGPPPPKPAERHKDPSGLVLPVGYKLGPNGLD 60
 DB 1 MAADGYLPDMLBTLSEGIQWMLKPGPPPPKPAERHKDPSGLVLPVGYKLGPNGLD 60
 QY 61 KGEVNEADAALHDKAYDQLSDGNDPILKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
 DB 61 KGEVNEADAALHDKAYDQLSDGNDPILKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
 QY 121 AKKRVLEPLGLVEPVTAPGKRPVHSPVEPDSSSGTGKAGQOPARKELNFGQTGDAD 180
 DB 121 AKKRVLEPLGLVEPVTAPGKRPVHSPVEPDSSSGTGKAGQOPARKELNFGQTGDAD 180
 QY 181 SVDPDQPLGQPPAPSGLTNTMTATGSGAPMADNNEGADGVNSGNHCDSTWMDRVI 240
 DB 181 SVDPDQPLGQPPAPSGLTNTMTATGSGAPMADNNEGADGVNSGNHCDSTWMDRVI 240
 QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYGYSTPMGYDFNPFCHFPDRWQRLI 300
 DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYGYSTPMGYDFNPFCHFPDRWQRLI 300
 QY 301 NNNWGPRLNFKLNIQVETQNDGTTIANNITSTVQVFTDSEYQLPYVLGSAHQG 360
 DB 301 NNNWGPRLNFKLNIQVETQNDGTTIANNITSTVQVFTDSEYQLPYVLGSAHQG 360
 QY 361 CLPPFPADVPMVQYGYLTINNGSQVGRSSFYCLEYFPPSQMLRTGNFTSYTTFEDVPF 420
 DB 361 CLPPFPADVPMVQYGYLTINNGSQVGRSSFYCLEYFPPSQMLRTGNFTSYTTFEDVPF 420
 QY 421 HSSVAHSQSLRLMNLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRDQSRNLWLP 480
 DB 421 HSSVAHSQSLRLMNLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRDQSRNLWLP 480

QY 481 PCYRQORVSKTSADNNSEYSWTGTATKYLHNGRDSLVPMPAMASHKDDDEKFFPQSGVL 540
 DB 481 PCYRQORVSKTSADNNSEYSWTGTATKYLHNGRDSLVPMPAMASHKDDDEKFFPQSGVL 540
 QY 541 IFGKGSGSEKTNVDIEKWMITDEEIRITNPNVATEQYGSVSTNLQGRNQAAATADVNTQGV 600
 DB 541 IFGKGSGSEKTNVDIEKWMITDEEIRITNPNVATEQYGSVSTNLQGRNQAAATADVNTQGV 600
 QY 601 LPMVMQDRDYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 DB 601 LPMVMQDRDYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 QY 661 FSAAKFASFTQYSTGVSEIWELOKENSKEWNEPEIQYTSYNNKSVNVDFVTDTNGVY 720
 DB 661 FSAAKFASFTQYSTGVSEIWELOKENSKEWNEPEIQYTSYNNKSVNVDFVTDTNGVY 720
 QY 721 SEPRPICTRYLTRNL 735
 DB 721 SEPRPICTRYLTRNL 735

RESULT 5
 ABB80231
 ID ABB80231 standard; protein; 735 AA.
 AC ABB80231;
 XX 20-NOV-2003 (first entry)
 DT AAV2 vp1 protein.
 DE Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;
 XX splice variant; transgene.
 KW Adeno associated virus serotype 2.
 OS WO2003052052-A2.
 PN 26-JUN-2003.
 PD 12-NOV-2002; 2002WO-US033631.
 XX 17-DEC-2001; 2001US-0341150P.
 PR 05-JUN-2002; 2002US-0386132P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA Gao G, Wilson JM, Alvira M;
 PI WPI; 2003-523523/49.
 DR New isolated adeno-associated virus (AAV) comprising an AAV9 capsid,
 XX useful for preparing a medicament for delivering a transgene to a cell.
 PT Disclosure; Fig 2; 76pp; English.

The sequences given in ABB80231-34 represent vp1 proteins derived from various adeno-associated virus (AAV) serotypes. These sequences were used in the scope of the invention for comparison with the cap protein derived from AAV serotype 9. The AAV capsid comprises three proteins vp1, vp2 and vp3, which are alternative splice variants. The AAV or the nucleic acid molecule is useful for preparing a medicament for delivering a transgene to a cell

Query Match 100.0%; Score 3994; DB 6; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAADGYLPDMLBTLSEGIQWMLKPGPPPPKPAERHKDPSGLVLPVGYKLGPNGLD 60
 DB 1 MAADGYLPDMLBTLSEGIQWMLKPGPPPPKPAERHKDPSGLVLPVGYKLGPNGLD 60

QY 661 FSAAKFASFTQYSTGQSVIEIWELOKENSKRWNPEIQTSYNKSNVNDFTVDINGVY 720
 Db 661 FSAAKFASFTQYSTGQSVIEIWELOKENSKRWNPEIQTSYNKSNVNDFTVDINGVY 720
 QY 721 SEPRPIGTRYLTRNL 735
 Db 721 SEPRPIGTRYLTRNL 735

RESULT 7
 ABR62761
 ID ABR62761 standard; protein; 735 AA.
 AC ABR62761;
 XX 06-NOV-2003 (first entry)
 DT
 XX Adeno associated virus 2 capsid protein vpl.
 DE
 XX AAV; AAV2; capsid; vector; gene therapy; antisense therapy; vaccine.
 KW Adeno associated virus.
 OS
 XX Key Location/Qualifiers
 FH Region 1. .184
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 24. .42
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 25. .28
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 81. .85
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 133. .165
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 134. .165
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 137. .143
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 146. .152
 FT /note= "hypervariable region, specifically referred to in
 FT Claim 4"
 FT Region 154. .156
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 171. .173
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 182. .187
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 185. .198
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 185. .198
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 194. .208
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 199. .259
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 260. .273
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"

FT Region 261. .274
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 262. .274
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 262. .264
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 263. .266
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 274. .446
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 381. .383
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 383. .385
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 413. .417
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 447. .477
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 449. .478
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 450. .474
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 451. .475
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 490. .495
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 491. .496
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 494. .525
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 495. .602
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 500. .504
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 501. .505
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 514. .522
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 533. .554
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 534. .571
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 534. .555
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 581. .601
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 581. .594
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 583. .596

XX PS 12-NOV-2002; 2002EP-00257826.
 XX PR 13-NOV-2001; 2001US-0350607P.
 XX PR 17-DEC-2001; 2001US-0341117P.
 XX PR 01-MAY-2002; 2002US-0377066P.
 XX PR 05-JUN-2002; 2002US-0386675P.
 XX PA (UYPE-) UNIV PENNSYLVANIA.
 XX PI Gao G, Wilson JM, Alvira M;
 XX DR WPI; 2003-450984/43.
 XX PS
 XX PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
 XX PT preventing or treating hyperproliferative or autoimmune diseases,
 XX PT comprises subjecting a sample having a DNA to amplification via
 XX PT polymerase chain reaction.
 XX PS
 XX PS Disclosure; SEQ ID NO 70; 419pp; English.
 XX PS
 XX CC The invention relates to a novel method for detecting adeno-associated
 XX CC virus (AAV) sequences in a sample, which comprises subjecting a sample
 XX CC containing a DNA to amplification via a polymerase chain reaction (PCR).
 XX CC The AAV sequence have the following activities: cytostatic,
 XX CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,
 XX CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
 XX CC sequence can be used in gene therapy or as part of a vaccine to treat
 XX CC disorders. The method is useful in detecting and/or identifying AAV
 XX CC sequences and isolating novel sequences that are identified. The
 XX CC sequences may be used e.g. for preventing or treating hyperproliferative
 XX CC conditions such as cancers and psoriasis, and other autoimmune diseases
 XX CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
 XX CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
 XX CC AAV related protein sequence of the invention.
 XX PS
 XX PS Sequence 735 AA;
 XX PS
 XX PS Query Match 100.0%; Score 3994; DB 7; Length 735;
 XX PS Best Local Similarity 100.0%; Pred. No. 0;
 XX PS Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX PS
 XX QY 1 MAADGYLDDWLEDTLSGIRQWKLKPGPPPPKPAERHKODSRLGLVLPYKYLGPFGNGLD 60
 XX DB 1 MAADGYLDDWLEDTLSGIRQWKLKPGPPPPKPAERHKODSRLGLVLPYKYLGPFGNGLD 60
 XX QY 61 KGEFVNEADAALHDXAYDQLSDGNPVLKYNHADAERFQRLKEDTSFGNIGRAVFG 120
 XX DB 61 KGEFVNEADAALHDXAYDQLSDGNPVLKYNHADAERFQRLKEDTSFGNIGRAVFG 120
 XX QY 121 AKKRVLPGLVPEPVKTAPEKKRPVSHSPVEPDSSSGTGKAGQAPARKLNFGQTGDAD 180
 XX DB 121 AKKRVLPGLVPEPVKTAPEKKRPVSHSPVEPDSSSGTGKAGQAPARKLNFGQTGDAD 180
 XX QY 181 SVDPDQPLGPPAPSGGLTNTWATGSGAPADNNEGADGVNSSGNWHCDSTWMDRVI 240
 XX DB 181 SVDPDQPLGPPAPSGGLTNTWATGSGAPADNNEGADGVNSSGNWHCDSTWMDRVI 240
 XX QY 241 TTSTRTWALPTYNHLYKQISSQGSASNDNHYFGYSTPMGYFDNRFHCFSPDRDWRLLI 300
 XX DB 241 TTSTRTWALPTYNHLYKQISSQGSASNDNHYFGYSTPMGYFDNRFHCFSPDRDWRLLI 300
 XX QY 301 NNNWGFPPKLNPKLFNIQVEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
 XX DB 301 NNNWGFPPKLNPKLFNIQVEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
 XX QY 361 CLPPFPADVFWPQYGYLTILNNGSQAVGRSSFFCLEYFSPQMLRTGNFTFSTFEDVDF 420
 XX DB 361 CLPPFPADVFWPQYGYLTILNNGSQAVGRSSFFCLEYFSPQMLRTGNFTFSTFEDVDF 420
 XX QY 421 HSSYAHSQSLDRLMNLPLIDLYLGRNTNTPSGTTTQSRLLQFSQAGASDIRQSRNWLFG 480
 XX DB 421 HSSYAHSQSLDRLMNLPLIDLYLGRNTNTPSGTTTQSRLLQFSQAGASDIRQSRNWLFG 480

481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDBEKFPQSGVL 540
 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDBEKFPQSGVL 540
 541 IFGQSGSEKTNVDIEKVMITDEBEIRTNPNVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
 541 IFGQSGSEKTNVDIEKVMITDEBEIRTNPNVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
 601 LPGMWODRDVYLGQPIWAKIPHTDGHFHPSPLMGGFGLKHPPOILLIKNTPVANPSTT 660
 601 LPGMWODRDVYLGQPIWAKIPHTDGHFHPSPLMGGFGLKHPPOILLIKNTPVANPSTT 660
 661 FSAAKFASFTQYSTQGVSVIEIWEIWKQENSKENWPEIQYTSNYSKSVNVDFTVDTNGVY 720
 661 FSAAKFASFTQYSTQGVSVIEIWEIWKQENSKENWPEIQYTSNYSKSVNVDFTVDTNGVY 720
 721 SEPRPIGTRVLTNRL 735
 721 SEPRPIGTRVLTNRL 735

RESULT 9
 AAB50326
 ID AAB50326 standard; protein; 734 AA.
 XX AAB50326;
 AC AAB50326;
 XX 09-MAR-2001 (first entry)
 DT
 XX Adeno-associated virus capsid protein sequence.
 DE
 XX Adeno-associated virus; AAV; capsid; virus binding inhibition;
 KW competitive inhibitor.
 KW
 XX Mastadenovirus.
 OS Synthetic.
 OS
 XX WO200073316-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US014466.
 PF
 XX 28-MAY-1999; 99US-00321589.
 PR
 XX (CELL-) CELL GENESYS INC.
 PA
 XX Patel S, McArthur J;
 PI
 XX WPI; 2001-061507/07.
 XX
 XX New polypeptide portion of a virus protein binding to an antibody
 PT specific for the virus useful for inhibiting binding of the virus to a
 PT cell or for binding host antibody to provide a transient tolerant or non-
 PT responsive state.
 PT
 XX Disclosure; Fig 2; 33pp; English.
 PS
 XX The present sequence is given in a specification relating to polypeptide
 CC portions of a virus protein or its derivative, that bind to an antibody
 CC specific for the virus or inhibit binding of the virus to a cell. The
 CC polypeptides are used in inhibiting the binding of viruses to cells of a
 CC host. Oligopeptides that inhibit binding of virus to receptor can be used
 CC as competitive inhibitors to release bound virus in an adsorption-type
 CC assay, and if an antibody was used as an immunoadsorbent, the
 CC oligopeptide could be used to elute bound virus from a solid support
 CC which virus antibody is immobilised. These oligopeptides may further be
 CC used to bind to host antibody to provide a transient tolerant or non-
 CC responsive state
 CC
 XX Sequence 734 AA;
 SQ

```
Query Match          99.3%; Score 3967.5; DB 4; Length 734;
Best Local Similarity 99.7%; Pred. No. 3.5e-317;
Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFQERLKEDTSFGNGLGRAVQ 120
Db 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFQERLKEDTSFGNGLGRAVQ 120

QY 121 AKKRVLPLGLVEEPVKTAPGKRKPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180
Db 121 AKKRVLPLGLVEEPVKTAPGKRKPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLGQPPAASGLGTNTMATGSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAASGLGTNTMATGSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240

QY 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300
Db 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300

QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360

QY 361 CLPPPPADVFMVPOQYGYLTLNNGSOAVGRSSFYCLEYFPSONLRTGNNFTSYTFEDVPP 420
Db 361 CLPPPPADVFMVPOQYGYLTLNNGSOAVGRSSFYCLEYFPSONLRTGNNFTSYTFEDVPP 420

QY 421 HSSYAHQSOLDRLMNLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
Db 421 HSSYAHQSOLDRLMNLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480

QY 481 PCYRQORVSKTSADNNSEYSWGATKYHLNGRDSLNVNPGPAMASHKODEEKFPPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSWGATKYHLNGRDSLNVNPGPAMASHKODEEKFPPQSGVL 540

QY 541 IFGKGSSEKTNVDIEKWMITDEEIRTTNPVATEQYGVSTNLQRNQAAATDVNTQGV 600
Db 541 IFGKGSSEKTNVDIEKWMITDEEIRTTN-VATEQYGVSTNLQRNQAAATDVNTQGV 599

QY 601 LPMGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPHPPQILIKNTVPVANPSTT 660
Db 601 LPMGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPHPPQILIKNTVPVANPSTT 659

QY 661 FSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYSNVNVDFTVDNGVY 720
Db 661 FSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYSNVNVDFTVDNGVY 719

QY 721 SEPRDGTGRYLRNL 735
Db 720 SEPRDGTGRYLRNL 734

RESULT 10
ID ADE76568
XX ADE76568 standard; protein; 735 AA.
XX ADE76568;
XX
XX 29-JAN-2004 (first entry)
XX
DE Adeno-associated virus (AAV) related protein, SEQ ID No 67.
XX
KW adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease.
```

```
XX Unidentified.
OS EP1310571-A2.
XX 14-MAY-2003.
XX 12-NOV-2002; 2002EP-00257826.
XX 13-NOV-2001; 2001US-0350607P.
XX 17-DEC-2001; 2001US-0341117P.
XX 01-MAY-2002; 2002US-0377066P.
XX 05-JUN-2002; 2002US-0386675P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Gao G, Wilson JM, Alvira M;
XX WPI; 2003-450984/43.
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
XX preventing or treating hyperproliferative or autoimmune diseases,
XX comprises subjecting a sample having a DNA to amplification via
XX polymerase chain reaction.
XX
XX Claim 24; SEQ ID NO 67; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
XX virus (AAV) sequences in a sample, which comprises subjecting a sample
XX containing a DNA to amplification via a polymerase chain reaction (PCR).
XX The AAV sequence have the following activities: cytostatic,
XX antipsoriatic, antirheumatic, antithyroid, dermatological, and antinflammatory. The AAV
XX sequence can be used in gene therapy or as part of a vaccine to treat
XX disorders. The method is useful in detecting and/or identifying AAV
XX sequences and isolating novel sequences that are identified. The
XX sequences may be used e.g. for preventing or treating hyperproliferative
XX conditions such as cancers and psoriasis, and other autoimmune diseases
XX like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
XX thyroiditis, scleroderma or Crohn's disease. This sequence represents an
XX AAV related protein sequence of the invention.
XX
XX Sequence 735 AA;
XX
Query Match          90.7%; Score 3623; DB 7; Length 735;
Best Local Similarity 89.1%; Pred. No. 7.8e-289;
Matches 655; Conservative 38; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60

QY 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFQERLKEDTSFGNGLGRAVQ 120
Db 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFQERLKEDTSFGNGLGRAVQ 120

QY 121 AKKRVLPLGLVEEPVKTAPGKRKPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180
Db 121 AKKRVLPLGLVEEPVKTAPGKRKPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLGQPPAASGLGTNTMATGSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAASGLGTNTMATGSGAPMADNNEGADGVGNSGNHCDSTWMDRVI 240

QY 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300
Db 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFCHFSPRDWQRLI 300

QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360

QY 361 CLPPPPADVFMVPOQYGYLTLNNGSOAVGRSSFYCLEYFPSONLRTGNNFTSYTFEDVPP 420
Db 361 CLPPPPADVFMVPOQYGYLTLNNGSOAVGRSSFYCLEYFPSONLRTGNNFTSYTFEDVPP 420
```

361 CLPPFPADVPMIQYGVLTLLNNGSQVAGRSSFFCLYFFSQMLRTGNNTFTSYTDEDVVF 420
 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
 481 PCYRQQRVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKDDSEKFFPQSGVL 540
 481 PSYRQQRMEKTDANNNSSEFAWTAATKYLNGRNSLVNPGPAMASHKDDSEKFFPQSGVL 540
 541 IFGKQSEKTNVDIEKVMITDEEERTNTPVATEQYGVSTNLQRNGRAATADVNTQGV 600
 541 IFGKQGTGTNVDIESVLITDEEERTNTPVATEQYGVSTNLQRNGRAATADVNTQGV 600
 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 661 FSAAKFASFIQYSGVSVIEWELQKENSKRWNPEIQYTSNKNKSVNVDTVDANGVY 720
 661 FPGKFAFIQYSGVSVIEWELQKENSKRWNPEIQYTSNKNKSVNVDTVDANGVY 720
 721 SEPRPIGTRYLTRNL 735
 721 SEPRPIGTRYLTRNL 735

RESULT 11
 ADE76570
 ID ADE76570 standard; protein; 735 AA.
 XX ADE76570;
 AC ADE76570;
 XX 29-JAN-2004 (first entry)
 XX Adeno-associated virus (AAV) related protein, SEQ ID No 69.
 XX adeno-associated virus; AAV; cytostatic; antipariatic; antirheumatic;
 KW antirheumatic; neuroprotective; antidiabetic; antithyroid;
 KW dermatological; antinflammatory; gene therapy; vaccine;
 KW hyperproliferative; cancer; psoriasis; autoimmune disease;
 KW rheumatoid arthritis; multiple sclerosis; diabetes;
 KW autoimmune thyroiditis; scleroderma; Crohn's disease.
 XX Unidentified.
 OS
 XX
 XX EP1310571-A2.
 XX 14-MAY-2003.
 XX 12-NOV-2002; 2002EP-00257826.
 XX 13-NOV-2001; 2001US-0350607P.
 PR 17-DEC-2001; 2001US-0341117P.
 PR 01-MAY-2002; 2002US-0370656P.
 PR 05-JUN-2002; 2002US-0386675P.
 XX (UTYPE-) UNIV PENNSYLVANIA.
 XX
 XX Gao G, Wilson JM, Alvira M;
 XX WPI; 2003-450984/43.
 XX
 XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
 PT preventing or treating hyperproliferative or autoimmune diseases,
 FT comprises subjecting a sample having a DNA to amplification via
 PT polymerase chain reaction.
 XX
 XX Claim 24; SEQ ID NO 69; 419pp; English.
 XX
 XX The invention relates to a novel method for detecting adeno-associated
 CC virus (AAV) sequences in a sample, which comprises subjecting a sample

CC containing a DNA to amplification via a polymerase chain reaction (PCR).
 CC The AAV sequence have the following activities: cytostatic,
 CC antipariatic, antirheumatic, antiarthritic, neuroprotective,
 CC antidiabetic, antirheumatoid, dermatological, and antinflammatory. The AAV
 CC sequence can be used in gene therapy or as part of a vaccine to treat
 CC disorders. The method is useful in detecting and/or identifying AAV
 CC sequences and isolating novel sequences that are identified. The
 CC sequences may be used e.g. for preventing or treating hyperproliferative
 CC conditions such as cancers and psoriasis, and other autoimmune diseases
 CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
 CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
 CC AAV related protein sequence of the invention.
 XX
 XX Sequence 735 AA;

Query Match 90.7%; Score 3622; DB 7; Length 735;
 Best Local Similarity 89.0%; Pred. No. 9.4e-289;
 Matches 654; Conservative 39; Mismatches 42; Indels 0; Gaps 0;
 QY 1 MAADGYLPDLWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPENGLD 60
 DB 1 MAADGYLPDLWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPENGLD 60
 QY 61 KGEVNEADAALAHDKAYDQLDSDGNPYLYKNHADAEFQERLKEDETSFGNGLGRAVQ 120
 DB 61 KGEVNEADAALAHDKAYDQLDSDGNPYLYKNHADAEFQERLKEDETSFGNGLGRAVQ 120
 QY 121 AKKRVLEPLGLVEEVPKTAPEKRPVVEHSPVPSDSSGTGKAGQQAPKRLNFGQTDAD 180
 DB 121 AKKRVLEPLGLVEEVPKTAPEKRPVVEHSPVPSDSSGTGKAGQQAPKRLNFGQTDAD 180
 QY 181 SVDPDQPLQGPAPPSGLGTNTMATSGAPMADNNEGADGVNNSGNHCHDSTMWMDRVI 240
 DB 181 SVDPDQPLQGPAPPSGLGTNTMATSGAPMADNNEGADGVNNSGNHCHDSTMWMDRVI 240
 QY 241 TTSTRTWALPTNNHLYKQISSQSGASNDHVFYGTSTPMGYDFNRFHCHESPRDWORLI 300
 DB 241 TTSTRTWALPTNNHLYKQISSQSGATNDHVFYGTSTPMGYDFNRFHCHESPRDWORLI 300
 QY 301 NNNWGRPKRLNPKLFNIOVKEVTQNDGTTTIANLLTSTVQVFTDSEYQLPYVLSAHQ 360
 DB 301 NNNWGRPKRLNPKLFNIOVKEVTQNDGTTTIANLLTSTVQVFTDSEYQLPYVLSAHQ 360
 QY 361 CLPPFPADVPMIQYGVLTLLNNGSQVAGRSSFFCLYFFSQMLRTGNNTFTSYTDEDVVF 420
 DB 361 CLPPFPADVPMIQYGVLTLLNNGSQVAGRSSFFCLYFFSQMLRTGNNTFTSYTDEDVVF 420
 QY 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
 DB 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
 QY 481 PCYRQQRVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKDDSEKFFPQSGVL 540
 DB 481 PSYRQQRMEKTDANNNSSEFAWTAATKYLNGRNSLVNPGPAMASHKDDSEKFFPQSGVL 540
 QY 541 IFGKQSEKTNVDIEKVMITDEEERTNTPVATEQYGVSTNLQRNGRAATADVNTQGV 600
 DB 541 IFGKQGTGTNVDIESVLITDEEERTNTPVATEQYGVSTNLQRNGRAATADVNTQGV 600
 QY 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 DB 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 QY 661 FSAAKFASFIQYSGVSVIEWELQKENSKRWNPEIQYTSNKNKSVNVDTVDANGVY 720
 DB 661 FPGKFAFIQYSGVSVIEWELQKENSKRWNPEIQYTSNKNKSVNVDTVDANGVY 720
 QY 721 SEPRPIGTRYLTRNL 735
 DB 721 SEPRPIGTRYLTRNL 735

AD E76569	standard; protein; 735 AA.
AD E76569	
29-JAN-2004	(first entry)
Adeno-associated virus (AAV) related protein, SEQ ID No 69.	
adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;	
antiarthritic; neuroprotective; antidiabetic; antichyroid;	
dermatological; antinflammatory; gene therapy; vaccine;	
hyperproliferative; cancer; psoriasis; autoimmune disease;	
rheumatoid arthritis; multiple sclerosis; diabetes;	
autoimmune thyroiditis; scleroderma; Crohn's disease.	
Unidentified.	
EP1310571-A2.	
14-MAY-2003.	
12-NOV-2002; 2002EP-00257826.	
13-NOV-2001; 2001US-0350607P.	
17-DEC-2001; 2001US-0341117P.	
01-MAY-2002; 2002US-0377066P.	
05-JUN-2002; 2002US-0386675P.	
(UNPE-) UNIV PENNSYLVANIA.	
Gao G, Wilson JM, Alvira M;	
WPI; 2003-450984/43.	
Detecting adeno-associated virus sequences in a sample, useful for e.g.	
preventing or treating hyperproliferative or autoimmune diseases,	
comprises subjecting a sample having a DNA to amplification via	
polymerase chain reaction.	
Claim 24; SEQ ID NO 68; 419pp; English.	
The invention relates to a novel method for detecting adeno-associated	
virus (AAV) sequences in a sample, which comprises subjecting a sample	
containing a DNA to amplification via a polymerase chain reaction (PCR).	
The AAV sequence have the following activities: cytostatic,	
antipsoriatic, antirheumatic, antiarthritic, neuroprotective,	
antidiabetic, antichyroid, dermatological, and antinflammatory. The AAV	
sequence can be used in gene therapy or as part of a vaccine to treat	
disorders. The method is useful in detecting and/or identifying AAV	
sequences and isolating novel sequences that are identified. The	
sequences may be used e.g. for preventing or treating hyperproliferative	
conditions such as cancers and psoriasis, and other autoimmune diseases	
like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune	
thyroiditis, scleroderma or Crohn's disease. This sequence represents an	
AAV related protein sequence of the invention.	
Sequence 735 AA;	
Query Match	90.6%; Score 3617; DB 7; Length 735;
Best Local Similarity	88.8%; Pred. No. 2.4e-288;
Matches 653; Conservative 41; Mismatches 41; Indels 0; Gaps 0;	
1 MAADGYPDLWEDTISEGIRQWKKLPGPPPPKPAERKDDSRGLVPGYKYLGPNGLD 60	
1 MAADGYPDLWEDTISEGIRQWKKLPGPPPPKPAERKDDSRGLVPGYKYLGPNGLD 60	
61 KGEYPNEADAALHDKAYDROLDSGDNPKYKNADAEFQRLKEDTSFGNIGRAVFO 120	
61 KGEYPNEADAALHDKAYDROLDSGDNPKYKNADAEFQRLKEDTSFGNIGRAVFO 120	
121 AKKRVLEPLGLVEEPAVKTA PGKKRPIEQSPAPEDSSSGIGESGQQA KAKRLNFGQTGDE 180	

Qy	121	AKKRVLEPLGLVEE	PVKTA	PGKKRP	VEHS	PVEP	DS	SSG	TG	KAG	QAP	KRL	NF	Q	TG	D	180
Db	121	AKKRILEPLGLVEE	AAAKTA	PGKKRP	VDOS	PQEP	DS	SSG	VG	KG	QK	P	KRL	NF	Q	TG	180
Qy	181	SVDPQPLGQCPA	PSGLG	TNTMA	TSG	SAP	MAD	NNE	GAD	VG	NS	SG	NH	CD	ST	W	240
Db	181	SVDPQPLGEP	PAAPT	SLG	SN	TMA	SGG	PAD	NNE	GAD	VG	NS	SG	NH	CD	S	240
Qy	241	TTSTRTWALPT	VNNHLY	KQI	SSQ	S	GA	SN	D	NH	YF	GY	ST	P	M	G	300
Db	241	TTSTRTWALPT	VNNHLY	KQI	SSQ	S	GA	SN	D	NH	YF	GY	ST	P	M	G	300
Qy	301	NNNGFRPKRL	NFKL	FN	I	QV	K	E	V	T	O	N	D	G	T	T	360
Db	301	NNNGFRPKKL	SFKL	FN	I	QV	K	E	V	T	O	N	D	G	T	T	360
Qy	361	CLPFP	PAD	V	M	V	P	O	Y	G	I	L	T	L	N	G	420
Db	361	CLPFP	PAD	V	M	V	P	O	Y	G	I	L	T	L	N	G	420
Qy	421	HSSVAHSQ	SLDR	L	N	P	L	I	D	Q	Y	L	L	Y	L	S	479
Db	421	HSSVAHSQ	SLDR	L	N	P	L	I	D	Q	Y	L	L	Y	L	S	479
Qy	480	GPCYRQ	R	V	K	S	T	A	D	N	N	S	E	S	W	T	539
Db	481	GPCYRQ	R	S	T	A	D	N	N	S	N	F	P	T	A	S	540
Qy	540	LI	FC	K	O	G	S	E	K	T	N	D	I	E	K	Y	599
Db	541	LI	FC	K	E	G	T	A	S	N	A	E	L	D	N	W	600
Qy	600	VLP	G	M	V	Q	D	R	V	Y	L	G	P	I	W	A	659
Db	601	ALP	G	M													

[illegible]

QY 720 YSEPRPIGTRYLTRNL 735
|||
Db 721 YSEPRPIGTRYLTRNL 736
|||

Search completed: May 19, 2005, 13:30:02
Job time : 169 secs